

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 2411.66 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

**Title:** US-09-899-276C-1

Perfect score:

Sequence: 1 taggaaaattataggatcat.....cctggaatcatccatttaa 600

Scoring table: IDENTITY NUC

Scoring table: IDENTIFY\_NOC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs. 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%  
Maximum Match 100%

Maximum MACCH 100%  
Listing first 45 summaries

**Database :**

Database : GenEmbl : ★

1: qb ba:\*

2: gb\_htg:★

```
3: gb_in:*
```

4: gb\_om: \*

5: gb\_ov:†

```
6: gb_pat:*
```

7: gb\_ph:\*

```
8: gb_pl:*
```

9: gb\_pr:\*

```
10: gb_ro:*
```

11: gb\_8ts:\*

```
12: gb_sy:*
13: gb_sy:*
```

```
13: gb_un:*
14: gb_un:*
```

14: gb\_v1:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		DB	ID	Description	
		Match	Length				
1	600	100.0	600	6	AX343328	Sequence	
2	600	100.0	9174	9	AF519531	Homo sapi	
3	600	100.0	11793	9	HS189333	Homo sapien	
c	4	600	100.0	147416	9	AC005549	Homo sapi
	5	121	20.2	2243	9	HSJEPJR	H. sapiens g
	6	71.6	11.9	242335	2	AC132303	Rattus no
7	71.6	11.9	245808	2	AC114440	Rattus no	
c	8	70.6	11.8	7218	6	I66494	Sequence 14
	9	60.4	10.1	160869	2	AC022239	Mus muscu
c	10	60.4	10.1	222121	10	AC012594	Mus muscu
c	11	60.4	10.1	223726	10	AL626807	Mouse DNR
	12	60.4	10.1	240425	10	CNS07YOT	AL713839
	13	55.8	9.3	180944	2	AC073824	Mus muscu
c	14	47.4	7.9	259772	2	AC106943	Rattus no
	15	44.6	7.4	239429	2	AC128200	Rattus no
c	16	44.6	7.4	243299	2	AC129672	Rattus no
c	17	44.6	7.4	265323	2	AC095987	Rattus no
	18	44.2	7.4	175254	2	AC127001	Rattus no
19	44.2	7.4	272717	2	AC098469	Rattus no	

## ALIGNMENTS

RESULT 1	AX343328	600 bp	DNA	linear	PAT 01-FEB-2002
LOCUS	Sequence 1 from Patent EP1170372.				
DEFINITION	AX343328				
ACCESSION	AX343328				
VERSION	AX343328.1	GI:18491678			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 Roesl, F., Soto, U., Coy, J., Finzer, P., Delius, H., Poustka, A., zur Hausen, H. and Patzelt, A.				
TITLE	Regulatory sequences of the human mcp-1 gene				
JOURNAL	Patent: EP 1170372-A 1 09-JAN-2002;				
	DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS (DE)				
FEATURES	Location/Qualifiers				
source	1..600				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%;	Score 600;	DB 6;	Length 600;	
Best Local Similarity	100.0%;	Pred. No. 1.7e-142;			
Matches	600;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TAGGAAATTTATAGGATCATTAAGACGAGGAGGAGGAGAGCTGGGAGCAAAATACCTGGAGG	60		
Db	1	TAGGAAATTTATAGGATCATTAAGACGAGGAGGAGGAGAGCTGGGAGCAAAATACCTGGAGG	60		
Qy	61	TAGAAATGGTGATGTGTGTACATCAAGCAGGGGAGAAAAACCAATGAACCAAGATGCCGAATT	120		
Db	61	TAGAAATGGTGATGTGTGTACATCAAGCAGGGGAGAAAAACCAATGAACCAAGATGCCGAATT	120		
Qy	121	CGGGCCCAACCAATGTCAGGGATGACATTTAGAAGGAGAGCTTGATGTCAGGGATTTC	180		
Db	121	CGGGCCCAACCAATGTCAGGGATGACATTTAGAAGGAGAGCTTGATGTCAGGGATTTC	180		
Qy	181	AATGTTAGGGGTGAAAAGTTACTACTCACTCTGTAGGTTAAAAAGGAACCGTTCAGGAATCT	240		

Db	181	AATGTTAGGTTGAAAGTTACTACTCAACTCTGTAGTTAAAGGAAACGTTGAGAACTCT	240		/rpt_family="MIR"
Qy	241	TCAGTCAATGAGAGGGATGTCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG	300		/rpt_type="dispersed"
Db	241	TCAGTCCAATGAGAGGGATGTCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG	300		/frequency="0.14"
Qy	301	TAACTTATAGATTTTATACATACACAGAGAAATACCGACTAGTGAGAACTATTGCCATG	360		/replace="g"
Db	301	TAACTTATAGATTTTATACATACACAGAGAAATACCGACTAGTGAGAACTATTGCCATG	360		/frequency="0.01"
Qy	361	GTCACAGCAGAGATGATGAAGCCTTAATATGAGACCCAAAGAGCGCAGCAATGAGAAATG	420		/replace="a"
Db	361	GTCACAGCAGAGATGATGAAGCCTTAATATGAGACCCAAAGAGCGCAGCAATGAGAAATG	420		/frequency="0.01"
Qy	421	AGCATCAGCGTGAATGCTGTCATGTTGAAATGAGGAGAAAGACCTGTGACTTTCAGA	480		/replace="t"
Db	421	AGCATCAGCGTGAATGCTGTCATGTTGAAATGAGGAGAAAGACCTGTGACTTTCAGA	480		/frequency="0.03"
Qy	481	TATGAAAACCTCATCTTTCACCCACATTTTAAAGGGGGAGCTTCCCTGAAACAGAAATGT	540		/replace="t"
Db	481	TATGAAAACCTCATCTTTCACCCACATTTTAAAGGGGGAGCTTCCCTGAAACAGAAATGT	540		/frequency="0.23"
Qy	541	GTTTCCCTCCATTTACTATACCCCATCCCAATCTCAGGCACCTGGAATCATCCATTAAA	600		/replace="a"
Db	541	GTTTCCCTCCATTTACTATACCCCATCCCAATCTCAGGCACCTGGAATCATCCATTAAA	600		/frequency="0.15"
RESULT 2					/replace="g"
AF519531		9174 bp	DNA	linear	/replace="a"
LOCUS	AF519531	Homo sapiens small inducible cytokine A2 (monocyte chemotactic			/frequency="0.04"
DEFINITION	AF519531	protein 1) (SCYA2) gene, complete cds.			/replace="g"
ACCESSION	AF519531.1	GI:21435976			/frequency="0.02"
VERSION					/replace="c"
KEYWORDS					/replace="t"
SOURCE					/frequency="0.08"
ORGANISM					/replace="a"
REFERENCE					/frequency="0.02"
AUTHORS					/replace="c"
TITLE					/frequency="0.13"
JOURNAL					/replace="t"
COMMENT					/frequency="0.02"
FEATURES					/replace="g"
source					/frequency="0.04"
variation					/replace="a"
variation					/frequency="0.01"
variation					/replace="a"
variation					/frequency="0.08"
variation					/replace="t"
variation					/frequency="0.03"
variation					/replace="a"
variation					/frequency="0.14"
variation					/replace="c"
variation					/rpt_family="MIR"
variation					/rpt_type="dispersed"
variation					/frequency="0.01"
variation					/replace="a"
variation					/frequency="0.08"
variation					/replace="t"
variation					/frequency="0.03"
variation					/replace="a"
variation					/frequency="0.14"
variation					/replace="c"
variation					/rpt_family="MIR"
variation					/rpt_type="dispersed"
variation					/frequency="0.01"
variation					/replace="a"
variation					/frequency="0.25"
variation					/replace="c"
variation					/gene="SCYA2"
variation					/gene="SCYA2"
variation					join(4065..4193,4990..5107,5490..5981)
variation					/product="small inducible cytokine A2 (monocyte
variation					chemotactic protein 1)"
variation					join(4118..4193,4990..5107,5490..5595)
variation					/gene="SCYA2"
repeat region					

```
/codon_start=1
/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
/protein_id="AAM54046.1"
/db_xref="GI:21435977"
/translation="MKVSAALLCLLLIAATFIQGLAQPDAINAPVTCYNFTNRKIS
VORLASRYRITSKCPKEAVIFKTIIVAKEICADPKQKWQDSMDHLDKOTQPKT"
4296..4409
repeat_region
/rpt_family="MIR"
/misc_feature
/rpt_type=dispersed
4359..4462
/genes="SCYA2"
/notes="Region not scanned for variation"
4748
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
repeat_region
4777..4814
/rpt_family="MIR"
/rpt_type=dispersed
4881
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
5018
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
5132
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
5660
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
6184
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
6518
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
6531
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
repeat_region
6837..6916
/rpt_family="L2"
/rpt_type=dispersed
7436
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
7605
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
7614
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
7659
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
7710
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
7843
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
variation
8072
/genes="SCYA2"
/replace="a"
/replace="c"
/replace="g"
/replace="t"
repeat_region
8118..8271
/rpt_family="MaLR"
/rpt_type=dispersed
8272..8734
repeat_region
```

```
/rpt_family="MaLR"
/rpt_type=dispersed
8597
/replace="a"
/replace="c"
/replace="g"
/replace="t"
repeat_region
8735..9033
/rpt_family="Alu"
/rpt_type=dispersed
8775
/replace="a"
/replace="c"
/replace="g"
/replace="t"
Query Match 100.0%; Score 600; DB 9; Length 9174;
Best Local Similarity 100.0%; Pred. No. 9.4e-143;
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAGGAAAATTATAGGATCATTTAAGAAAGGAGAGAGTGGGAGCAAAATACCTGGAGG 60
DB 6501 TAGGAAAATTATAGGATCATTTAAGAAAGGAGAGAGTGGGAGCAAAATACCTGGAGG 6560
QY 61 TAGAATGGTGATGTGTACATCAAGCAGGAGAAACCAATGAACCCAGATGCGAATT 120
DB 6561 TAGAATGGTGATGTGTACATCAAGCAGGAGAAACCAATGAACCCAGATGCGAATT 6620
QY 121 CGGGCCCCACACCAATGTCAAGGGATGACAAATTAGAAAGGAGAGTGTGAGTCAAGGGATTG 180
DB 6621 CGGGCCCCACACCAATGTCAAGGGATGACAAATTAGAAAGGAGAGTGTGAGTCAAGGGATTG 6680
QY 181 AATGTTAGGGTGAAGAGTTACTCAACTCTGTAGTTTAAAGGAAAGCTTGAGATCT 240
DB 6681 AATGTTAGGGTGAAGAGTTACTCAACTCTGTAGTTTAAAGGAAAGCTTGAGATCT 6740
QY 241 TCAGTCCCAATGAGGAGGATGCGCATGTTTAGAGATTACAGAGATAAGTTTCCAGGAAATG 300
DB 6741 TCAGTCCCAATGAGGAGGATGCGCATGTTTAGAGATTACAGAGATAAGTTTCCAGGAAATG 6800
QY 301 TAACTTATAGATTTTATACATACACAGAGAAATACGGACTAGTGAGAGCTATTGCCATG 360
DB 6801 TAACTTATAGATTTTATACATACACAGAGAAATACGGACTAGTGAGAGCTATTGCCATG 6860
QY 361 GTCCAAGCAAGAGATGATGAAGCCCTAAATATGAGCCAAAGAGCAGCAATGAAGAATG 420
DB 6861 GTCCAAGCAAGAGATGATGAAGCCCTAAATATGAGCCAAAGAGCAGCAATGAAGAATG 6920
QY 421 AGCCATGAGGGTGAAGTCTGCATGTTGTAATCGAGGAGAAACACCTGTCACTTCAGA 480
DB 6921 AGCCATGAGGGTGAAGTCTGCATGTTGTAATCGAGGAGAAACACCTGTCACTTCAGA 6980
QY 481 TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGCGAGCTTCCTCGAAACAGAAATG 540
DB 6981 TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGCGAGCTTCCTCGAAACAGAAATG 7040
QY 541 GTTTCCTCCATTACTATACCCCATCCCAATCTCAGGACCTGGAATCATCAATTTAAA 600
DB 7041 GTTTCCTCCATTACTATACCCCATCCCAATCTCAGGACCTGGAATCATCAATTTAAA 7100
RESULT 3
HSY18933
LOCUS
DEFINITION Homo sapiens MCP-1 gene and enhancer region.
ACCESSION Y18933
VERSION Y18933.1 GI:10933860
KEYWORDS MCP-1 gene; monocyte chemoattractant protein-1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Finzer,P., Soto,U., Delius,H., Patzelt,A., Coy,J.F., Pouetka,A., zur Hausen,H. and Rosl,F.
TITLE Differential transcriptional regulation of the monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and
```

non-tumorigenic HPV 18 positive cells: the role of the chromatin structure and AP-1 composition  
Oncogene 19 (29), 3235-3244 (2000)

JOURNAL  
MEDLINE

PUBMED  
10918580

REFERENCE  
2 (bases 1 to 11793)

AUTHORS  
Roel, F.

TITLE  
Direct Submission

JOURNAL  
Submitted (15-APR-1999) F. Roel, Applied Tumor Virology, DKFZ  
(German Cancer Research Institute), Im Neuenheimer Feld 506, 69120  
Heidelberg, FRG

FEATURES

Location/Qualifiers

1. .11793

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="17"

4906. .9073

/gene="MCP-1"

4906. .4915

/gene="MCP-1"

4935. .4944

/gene="MCP-1"

7513. .7517

/gene="MCP-1"

join(<7550. .7685,8468. .8585,8968. .>9073)

/gene="MCP-1"

<7550. .7685

/gene="MCP-1"

/number=1

join(7610. .7685,8468. .8585,8968. .9073)

/gene="MCP-1"

/codon\_start=1

/product="monocyte chemoattractant protein-1"

/protein\_id="CAC14049.1"

/db\_xref="GI:10933861"

/db\_xref="GOA:P13500"

/db\_xref="UniProt/Swiss-Prot:P13500"

/translations="MKVSAALLCLLIATFIQGLAQPDAINAPVTCYNTNRKIS  
VQLASVRRITSSKCPKEAVIFKTIIVAKEICADPKQKQVDSNDHLDKQTQPKT"

7686. .8467

/gene="MCP-1"

/number=1

8468. .8585

/gene="MCP-1"

/number=2

8586. .8967

/gene="MCP-1"

/number=2

8968. .>9073

/gene="MCP-1"

/number=3

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 600; DB 9; Length 11793;

Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TAGGAAATTTAGGATCATTTAAGAAAGGAGAGAGTGGGAGCAATACCTGGAGG 60

|||||

9979 TAGGAAATTTAGGATCATTTAAGAAAGGAGAGTGGGAGCAATACCTGGAGG 10038

|||||

61 TAGAATGGTGTGATGTGATCATCAACGACGGAGAAACCAATGAACGATCGGAATT 120

|||||

10039 TAGAATGGTGTGATGTGATCATCAACGACGGAGAAACCAATGAACGATCGGAATT 10098

|||||

121 CGGGCCACACCAATGTCAAGGGATGACAATTAGAAAGGAGAGGTGTGAGTCAAGGGATTG 180

|||||

10099 CGGGCCACACCAATGTCAAGGGATGACAATTAGAAAGGAGAGGTGTGAGTCAAGGGATTG 10158

|||||

181 AATGTTAGGGTGAAGATTTACTACTCACTCTGTAGGTTTAAAGGAAACGTTGAGATCT 240

|||||

10159 AATGTTAGGGTGAAGATTTACTACTCACTCTGTAGGTTTAAAGGAAACGTTGAGATCT 10218

241 TCAGTCCATAGGAGGAGTGTGCCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG 300  
|||||

10219 TCAGTCCATAGGAGGAGTGTGCCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG 10278

|||||

301 TAACTTTATAGATTTTATACATACACAGAGAAATACGACTAGTGAGAACTATTGCGCATG 360

|||||

10279 TAACTTTATAGATTTTATACATACACAGAGAAATACGACTAGTGAGAACTATTGCGCATG 10338

|||||

361 GTCCAAGCAAGAGATGATGAAGCCCTAAATATGGAGCCCAAGAGGAGGAGCAATGAAGAATG 420

|||||

10339 GTCCAAGCAAGAGATGATGAAGCCCTAAATATGGAGCCCAAGAGGAGGAGCAATGAAGAATG 10398

|||||

421 AGCCATGCAGGGTGAATGCTGTCATGTTGTAATATGAGGAGAAAGACCTGTGACTTCAGA 480

|||||

10399 AGCCATGCAGGGTGAATGCTGTCATGTTGTAATATGAGGAGAAAGACCTGTGACTTCAGA 10458

|||||

481 TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGAGCTTCCCTGAAACCAAGATGT 540

|||||

10459 TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGAGCTTCCCTGAAACCAAGATGT 10518

|||||

541 GTTTCCTTCATTTACTATACCCCATCCCAATCTCAGGCACCTGGAAATCATCCATTAAA 600

|||||

10519 GTTTCCTTCATTTACTATACCCCATCCCAATCTCAGGCACCTGGAAATCATCCATTAAA 10578

|||||

RESULT 4

AC005549/c 147416 bp DNA linear PRI 22-SEP-1998

LOCUS Homo sapiens chromosome 17, clone hRPK.215\_E\_13, complete sequence.

DEFINITION AC005549

ACCESSION AC005549.1 GI:3598724

VERSION HTG.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 147416)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 17, clone hRPK.215\_E\_13

Unpublished

2 (bases 1 to 147416)

Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,

Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,

Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M.,

Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E.,

Devon, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W.,

Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraghty, K., Grant, G.,

Hagos, B., Heaford, A., Herena, L., Horton, L., Howland, J. C.,

Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,

Macdonald, P., Marquis, N., McSwan, P., McGurk, A., McKernan, K.,

Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,

Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,

Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A.,

Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,

Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,

Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,

Ye, W. J., Zhao, J. and Zody, M.

Direct Submission

Submitted (27-AUG-1998) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 147416)

Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,

Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,

Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M.,

Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E.,

Devon, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W.,

Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraghty, K., Grant, G.,

Hagos, B., Heaford, A., Herena, L., Horton, L., Howland, J. C.,

Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,

Macdonald, P., Marquis, N., McSwan, P., McGurk, A., McKernan, K.,

Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,

Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,

Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P.,  
 Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,  
 Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,  
 Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,  
 Ye, W. J., Zhao, J. and Zody, M.

# TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 147416)

Birren, B., Fagan, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
 Anderson, M., Baker, J., Baldwin, J., Barna, N., Becker, R., Benn, J.,  
 Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M.,  
 Collins, S., Collins, A., Cooke, P., Corliss, D., Depayre, E.,  
 Devor, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W.,  
 Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G.,  
 Hages, B., Heaford, A., Herena, L., Horton, L., Howland, J. C.,  
 Jacotot, J., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,  
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
 Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,  
 Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,  
 Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P.,  
 Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,  
 Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,  
 Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,  
 Ye, W. J., Zhao, J. and Zody, M.

# TITLE

## JOURNAL

## COMMENT

Submitted (22-SEP-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 13, 1998 this sequence version replaced gi:3581743.

All repeats were identified using RepeatMasker: Smit, A.F.A. &

Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Only the first 147416 bases of this clone are being submitted.  
 Bases 145417-155040 overlap accession number AC004147 (WICGR  
 project L228). The first 2Kb of the overlapping region are  
 submitted to confirm overlap.

## FEATURES

## Source

1..147416  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="17"  
 /map="17"  
 /clone="hRPK 215 E 13"  
 /clone\_lib="RPCT-11 human BAC library"

2..69

/rpt\_family="BUR1"

490..695

/rpt\_family="LIME2"

1291..1390

/rpt\_family="MIR"

2980..3131

/rpt\_family="MER5A"

complement(3147..3241)

/rpt\_family="MIR"

complement(4228..4469)

/rpt\_family="MLT11"

complement(5430..5549)

/rpt\_family="FLAM\_C"

complement(5941..6051)

/rpt\_family="MER81"

6131..6250

/rpt\_family="(TA)n"

6406..6455

/rpt\_family="AT\_rich"

6574..6600

/rpt\_family="MLT1D"

6601..6680

/rpt\_family="AluJb"

6881..6935

/rpt\_family="MLT1D"

7119..7360

repeat\_region  
 7337..7636  
 /rpt\_family="MLT1D"  
 repeat\_region  
 7689..7888  
 /rpt\_family="LIME3"  
 repeat\_region  
 8023..8048  
 /rpt\_family="AT\_rich"  
 repeat\_region  
 8871..8986  
 /rpt\_family="MIR"  
 repeat\_region  
 10109..10258  
 /rpt\_family="MER5A"  
 complement(11031..11072)  
 /rpt\_family="MIR"  
 repeat\_region  
 11567..11836  
 /rpt\_family="L2"  
 repeat\_region  
 12501..12981  
 /rpt\_family="MER9"  
 complement(12983..13257)  
 /rpt\_family="AluJb"  
 repeat\_region  
 15248..15364  
 /rpt\_family="MIR"  
 complement(16483..16653)  
 /rpt\_family="MIR"  
 repeat\_region  
 16830..16949  
 /rpt\_family="MER5A"  
 repeat\_region  
 17203..17380  
 /rpt\_family="LTR16A"  
 repeat\_region  
 17393..18306  
 /rpt\_family="HERVL"  
 repeat\_region  
 18315..24307  
 /rpt\_family="L1PA2"  
 repeat\_region  
 24309..24335  
 /rpt\_family="AT\_rich"  
 repeat\_region  
 24513..25245  
 /rpt\_family="HERVL"  
 repeat\_region  
 25613..25758  
 /rpt\_family="HERVL"  
 repeat\_region  
 26135..26228  
 /rpt\_family="LTR16A"  
 complement(27814..27950)  
 /rpt\_family="MER94"  
 repeat\_region  
 28930..28968  
 /rpt\_family="AT\_rich"  
 repeat\_region  
 29960..30341  
 /rpt\_family="MLT11"  
 repeat\_region  
 32364..32581  
 /rpt\_family="MIR"  
 repeat\_region  
 32601..32902  
 /rpt\_family="AluSx"  
 complement(33593..39359)  
 /rpt\_family="L1PA2"  
 repeat\_region  
 39382..39414  
 /rpt\_family="A-rich"  
 complement(39920..40047)  
 /rpt\_family="MIR"  
 repeat\_region  
 40112..40151  
 /rpt\_family="AT\_rich"  
 complement(40152..40442)  
 /rpt\_family="AluJo"  
 complement(40810..40898)  
 /rpt\_family="MLT1J"  
 repeat\_region  
 41293..41469  
 /rpt\_family="L1PA6"  
 complement(41846..41956)  
 /rpt\_family="purine-rich"  
 complement(42486..42569)  
 /rpt\_family="MLT11"  
 repeat\_region  
 42979..43098  
 /rpt\_family="MER5A"  
 repeat\_region  
 43554..43983  
 /rpt\_family="LTR24"  
 repeat\_region  
 44781..44811  
 /rpt\_family="AT\_rich"

**TITLE**  
JOURNAL  
Submitted (16-MAY-1991) E. Schwarz, Albert Einstein College of Medicine, Dept of Microbiology & Immunology, Forchheimer Bldg Room 411, 1300 Morris Park Avenue, Bronx NY 10461, USA

**FEATURES**  
source  
location/Qualifiers  
1. .2243  
    /organism="Homo sapiens"  
    /mol\_type="genomic DNA"  
    /db\_xref="taxon:9606"  
    /tissue\_type="liver"  
    /clone\_lib="Okayama-Berg plasmid"  
    /dev\_stage="fetal"  
    <l..612  
        /number=1  
    /gene="gamma 1,hje"  
        613..1591  
    /exon  
        613..730  
    /gene="gamma 1,hje"  
    /product="JE protein"  
    /number=2  
    731..1112  
    /gene="gamma 1,hje"  
    /number=2  
    1113..1591  
    /exon  
        /gene="gamma 1,hje"  
    /product="JE protein"  
    /number=3

**ORIGIN**

Query Match	20.2%;	Score 121;	DB 9;	Length 2243;
Best Local Similarity	100.0%;	Pred. No. 3e-20;		
Matches 121;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY     1   TAGCAAAATTATAGGATCATTTAAGAAAGAGAGAGAGAGACCAATTAACCTGGAGG 60  
Db     2123 TAGCAAAATTATAGGATCATTTAAGAAAGAGAGAGAGAGACCAATTAACCTGGAGG 2182

QY     61   TAGCAATGGTGATGTGTGCATCAAGCAGGAGAGAAAACCAATGAACAGATGCCAATT 120  
Db     2183 TAGCAATGGTGATGTGTGCATCAAGCAGGAGAGAAAACCAATGAACAGATGCCAATT 2242

QY     121 C 121  
Db     2243 C 2243

**RESULT 6**  
AC123203  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-28P9, WORKING DRAFT SEQUENCE.  
ACCESSION  
AC123203  
VERSION  
HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
Rattus norvegicus  
SOURCE  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

**REFERENCE**  
**AUTHORS**  
1. (bases 1 to 242335)  
Murphy,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen.C.,Allen.H., Alsbrooks,S., Amin,A., Anguiano.D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Biankburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burn,P., Surenell,K., Calderon.E., Cardenas,V., Carter.K., Cavazos,I., Cesar,H., Chen,X., Chu.J., Chacko,J., Chavez,D., Chen,G., Chen,R., Coyle,M., Cree.A., D'Souza,L., Cleveland,C., Cockrell.R., Cox,C., Coyne,M., Crease,A., Dederich,D., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto.M., Eugene,C., Evans.C.A., Falls.T., Fan,G., Fernandez.S., Finley.M., Flagg,N., Forbes,L., Foster,M., Foster,P.,



and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
-----
Center project name: GSAK
Center clone name: CH230-12P3
-----
Summary Statistics
-----
Assembly program: Phrap; version 0.990329
Consensus quality: 226189 bases at least Q40
Consensus quality: 228217 bases at least Q30
Consensus quality: 229460 bases at least Q20
Consensus quality: 230674; sum-of-contrigs
Estimated insert size: 230674; sum-of-contrigs
-----

```

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* be preserved.  
\* 1 3901: contig of 3901 bp in length  
\* 3902 4001: gap of unknown length  
\* 3903 24500: contig of 241807 bp in length.

## FEATURES

```

source
1..245908
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-12193"

misc_feature
1..1014
/note="wgs contig"

misc_feature
4002..5165
/note="wgs contig"

misc_feature
5719..7604
/note="wgs contig"

```

## ORIGIN

	Query Match	11.9%;	Score 71.6;	DB 2;	Length 245808;
	Best Local Similarity	57.0%;	Pred.No. 4.7e-08;	Mismatches 129;	Gaps 3;
	Matches 203;	Conservative	0;	Indels	
QY	5	A A A A T T A T A G C A T C A T T A C A A A G G A G A A G G A A G T G G G G A C C A A A T A C C T G G A G G T A G A	64		
Db	130304	A A A A A A A C A G A A G C A T T T A G A A A - A C A A G A A A G G T G G G G G G A A T T A T C C A A A G G T A G A	130361		
QY	65	A A T G T G T A T G A T G T G T A C A T C A A G C A G G G A G A A A C C A A T G A A C C A G A T G C G A A T T C G G G	124		
Db	130362	A A T G A A G A T G T G T G T A T G T C A A A G G G G G A A G A - - - - - G A A C T A T G	130403		
QY	125	C C C A C A C C A A T G T C A A G G G A T G A C A A T T A G A A A G A A G G T T G A G T C A A G G G A T T T G A A T G	184		
Db	130404	C A G C G T G T A T G T C A A G A A T G G G G T T T T C A A A A G G A A G T G A A T C A C G G T T T T A A T G	130463		
QY	185	T T A G G G T G A A A G T T A C T A C T C A A C T C T G T A G G T T A A A A G C A A A C G T T G A G A A T C T T C A G	244		
Db	130464	T A A A G T G A G A A C A T G G T G G T C C C G G T G A C C T G A T T T T A A A G A A G C T G A A A C T C T T T G G	130523		

AC114440.4	GI:25138937
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.	
Rattus norvegicus (Norway rat)	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
REFERENCE	
AUTHORS	1 (bases 1 to 245808)
TITLE	Murzy, D., Maritz, Metzker, M.Lee., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalbechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buñay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J.J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.C., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escoto-Mora, Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, W., Guerra, W., Gregorogorski, P., Haaland, W., Hamil, C., Hamilton, C., Guevara, W., Gunaratne, P., Hawes, A., Henderson, N., Hernandez, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huliy, S., Hume, J., Idlebird, R., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kovac, C., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C., Kowic, C., Kraft, C.L., Lebow, H., Levian, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louisegh, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, N., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaokemeleh, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polozdexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sittler, C.D., Smajda, D., Snead, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlaczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 245808)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 245808)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Nov 20, 2002 this assembly version replaced gi:23195167. The sequence in this assembly is a combination of BAC based reads



Db	129077	AATAGGAAGGTGAGAACATGGTATTCCCTCGGACCTATATAATTTAAAGGATAAGTTGAAC	129018
QY	238	TCTTCAGTCCAATGA-----GGAGGATGTGCCATGTTTAGAGATTTCAGAGATAAGTTTCA	293
Db	129017	TCTTCGGTCCAATAAAGAGGAGGAGCGTGTTCATTCAGAGGCTCAGAGGTGAGCCTCA	128958
QY	294	G-GAAATGTAACCTTATATAGATTTTATACATACACAGAGAAATACGGACTAGTGAGAGCTA	352
Db	128957	GAAAATAATACATCCTGATGACTTCAGACAGACAGAAAGAGTGGAGACAAAGTGTGGTTG	128898
QY	353	TTGC 356	
Db	128897	TGGC 128894	
RESULT 10			
AC012294			
Mus musculus chromosome 11, clone RP23-328G11, complete sequence.			
AC012294			
AC012294.15 GI:19703299			
HTG.			
Mus musculus (house mouse)			
Mus musculus			
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.			
1 (bases 1 to 222121)			
Mus musculus chromosome 11, clone RP23-328G11			
2 (bases 1 to 222121)			
Unpublished			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

```

Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 222121)
Birren,B., Linton,L., Nusbbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukgafter,B., Brown,A., Camarato,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,A., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 25, 2002 this sequence version replaced gi:19683729.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3202
Center clone name: 328_G_11
-----
Location/Qualifiers
1. .222121
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-328G11"
/complement(1..315)
/rpt_family="MTD"
403..560
/rpt_family="MER109"
/complement(793..916)
/rpt_family="ORR1B-int"
1209..1404
/rpt_family="LIME"
1632..1714
/rpt_family="LIME"
/complement(2080..2434)
/rpt_family="LIMC4a"
2897..2935
/rpt_family="(TCTG)n"
2965..4300
/rpt_family="L1"
/complement(5229..6151)
/rpt_family="L1"
6159..6709
/rpt_family="Lx2"
6718..6747
/rpt_family="(CAAAA)n"
/complement(7561..7875)
/rpt_family="Lx7"

repeat_region      7876..7930
/rpt_family="A-rich"
complement(7939..8115)
/rpt_family="Lx7"
repeat_region
complement(8119..8409)
/rpt_family="L1"
8413..8451
/rpt_family="(TG)n"
8466..8658
/rpt_family="Lx5"
complement(8659..9053)
/rpt_family="MTA"
9054..9357
/rpt_family="Lx5"
9355..15501
/rpt_family="L1_MM"
15954..15979
/rpt_family="(CAAA)n"
16045..16191
/rpt_family="B1_MM"
16562..16585
/rpt_family="(CAAA)n"
16832..16931
/rpt_family="RSINE1"
16974..17049
/rpt_family="(CAAA)n"
18843..19008
/rpt_family="T-rich"
complement(19010..19022)
/rpt_family="B1_MM"
19023..19192
/rpt_family="TTCTC)n"
complement(19193..19331)
/rpt_family="B1_MM"
complement(19335..19654)
/rpt_family="MTD"
complement(19655..20059)
/rpt_family="RLTR14"
20212..20409
/rpt_family="MT2A"
20596..20641
/rpt_family="(CA)n"
24235..24267
/rpt_family="(CAAA)n"
24926..25139
/rpt_family="Lx5"
complement(25142..25721)
/rpt_family="L1"
25722..26248
/rpt_family="L1_MM"
26241..26342
/rpt_family="Lx5"
26655..26722
/rpt_family="(GA)n"
26903..27332
/rpt_family="RMERS"
27737..28052
/rpt_family="MT-int"
27809..27813
/Note="<30 qual SNGL region"
28254..28598
/rpt_family="MTD"
28828..28906
/rpt_family="MTC"
complement(28973..29651)
/rpt_family="LIM4"
complement(29778..29944)
/rpt_family="B4A"
complement(30238..30446)
/rpt_family="LIM4"
30463..30483

10.1%; Score 60.4; DB 10; Length 222121;
Query Match

```

REFERENCE  
AUTHORSTITLE  
JOURNAL

## COMMENT

FEATURES  
source

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

Tue Aug 9 17:01:24 2005

```

FEATURES
  source
    1. 223726
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="11"
    /clones="RP23-350G1"
    /clone_lib="RPCI-23"

ORIGIN
  Query Match      10.1%; Score 60.4; DB 10; Length 223726;
  Best Local Similarity 55.5%; Pred. No. 3.5e-05;
  Matches 202; Conservative 0; Mismatches 151; Indels 11; Gaps 4;

  1 TAGGAAAATTATAGGATCATTAAGAAAGGAGAGAGTGGGAGCAAAATACCTGGAGG 60
  |||||
  100274 TAGGAAAATAACACAAAAGCATTAAGGAAACCTGGTGGTGGGGAAATATCCAAAA 100215
  |||||
  61 TAGAATGTGTGATGTGTATCATCAAGCAGGAGAGAAACCAATGAACCGATGCGAATT 120
  |||||
  100214 TAGAATGAATATGGTGTGTATCAATTAACGGGAGAGCGGAGGAAACCATCCG---ATG 100158
  |||||
  121 CGGGCCACACCAATGTCAAGGGATGACATTTAGAAAGGAGGTTTGATCAAGGATTG 180
  |||||
  100157 GAGCTCATGTATATCAAGAGATGGGGATTTTCAAGAGGAGCTCAATACAGGGTTGG 100098
  |||||
  181 AATGTTAGGTGAAGATTTACTCAACT---CTGTAGGTTTAAAGGAAACGTTGAGAA 237
  |||||
  100097 AATAGGAAGGTGAGAACATGTTATCCCTCGGACCTATTAATTTAAAGGATAAGTTGAAAC 100038
  |||||
  238 TCTTCAGTCCAATGA----GGAGGATGTGCCATGTTTAGAGATTCAGAGATAAGTTTCA 293
  |||||
  100037 TCTTCGTCCTCAATAAAGAGGGAGGACGTTGTTTCATTCAGAGGCTCAGAGGTGACCTCA 99978
  |||||
  294 G-GAAATGTAATATATAGATTTTATACATACACAGAGAAATACGGACTAGTCAGAAAGCTA 352
  |||||
  99977 GAAATAATACATCTCTGATGACTTCAGACAGACAGAGAAAGTGGAGACAAGTGTATGGTTG 99918
  |||||
  353 TTGC 356
  |||||
  99917 TGGC 99914

RESULT 11
AL626807/c      223726 bp DNA linear ROD 05-APR-2002
LOCUS           Mouse DNA sequence from clone RP23-350G1 on chromosome 11, complete
DEFINITION      sequence.
ACCESSION       AL626807.7 GI:17221258
VERSION         AL626807.7
KEYWORDS        HTG.
SOURCE          Mus musculus (house mouse)
ORGANISM        Mus musculus (house mouse)

REFERENCE
  1. Whitehead, S.
  Direct Submission
  Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Nov 30, 2001 this sequence version replaced gi:17065774.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one plasmid subclone or more than one M13 subclone; and the
  assembly was confirmed by restriction digest. The following
  abbreviations are used to associate primary accession numbers given
  in the feature table with their source databases: Em: EMBL; Sw:
  SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
  database can be found at
  http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-350G1 is
  from the RPCI-23 Mouse PAC Library
  constructed by the group of Pieter de Jong.
  For further details see http://www.chori.org/bacpac/home.htm
  VECTOR: pBACE3.6.

```

```

FEATURES
  source
    1. 223726
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="11"
    /clones="RP23-350G1"
    /clone_lib="RPCI-23"

ORIGIN
  Query Match      10.1%; Score 60.4; DB 10; Length 223726;
  Best Local Similarity 55.5%; Pred. No. 3.5e-05;
  Matches 202; Conservative 0; Mismatches 151; Indels 11; Gaps 4;

  1 TAGGAAAATTATAGGATCATTAAGAAAGGAGAGAGTGGGAGCAAAATACCTGGAGG 60
  |||||
  9452 TAGGAAAATAACACAAAAGCATTAAGGAAACCTGGTGGTGGGGAAATATCCAAAA 9393
  |||||
  61 TAGAATGTGTGATGTGTATCATCAAGCAGGAGAGAAACCAATGAACCGATGCGAATT 120
  |||||
  9392 TAGAATGAATATGGTGTGTATCAATTAACGGGAGAGCGGAGGAAACCATCCG---ATG 9336
  |||||
  121 CGGGCCACACCAATGTCAAGGGATGACATTTAGAAAGGAGGTTTGATCAAGGATTG 180
  |||||
  9335 GAGCTCATGTATATCAAGAGATGGGGATTTTCAAGAGGAGCTCAATACAGGGTTGG 9276
  |||||
  181 AATGTTAGGTGAAGATTTACTCAACT---CTGTAGGTTTAAAGGAAACGTTGAGAA 237
  |||||
  9275 AATAGGAAGGTGAGAACATGTTATCCCTCGGACCTATTAATTTAAAGGATAAGTTGAAAC 9216
  |||||
  238 TCTTCAGTCCAATGA----GGAGGATGTGCCATGTTTAGAGATTCAGAGATAAGTTTCA 293
  |||||
  9215 TCTTCGTCCTCAATAAAGAGGGAGGACGTTGTTTCATTCAGAGGCTCAGAGGTGAGCTCA 9156
  |||||
  294 G-GAAATGTAATATATAGATTTTATACATACACAGAGAAATACGGACTAGTCAGAAAGCTA 352
  |||||
  9155 GAAATAATACATCTCTGATGACTTCAGACAGACAGAGAAAGTGGAGACAAGTGTATGGTTG 9096
  |||||
  353 TTGC 356
  |||||
  9095 TGGC 9092

RESULT 12
CNS07YOT      240425 bp DNA linear ROD 23-MAR-2002
LOCUS         Mus musculus chromosome 11 region in the Om locus area
DEFINITION    (D11Mit37-Scy6) clone 149H13 of library Caltech CITB-BAC from
              chromosome 11 of Mus musculus (mouse).
ACCESSION     AL713839.1 GI:19682818
VERSION       366M1-SP6; D11Bhm140; D11Mit354; D11Mit66; DDK syndrome; Mdeg; Ovum
KEYWORDS      mutant; Scy2,11.7,8,12,1.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
              1 (bases 1 to 240425)
              Genoscope.
REFERENCE     Direct Submission
AUTHORS      Submitted (20-MAR-2002) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              Web : www.genoscope.cns.fr)
COMMENT       IMPORTANT: This sequence is unfinished and does not necessarily
              represent the correct sequence.
              Cohen-Tannoudji M., Vandormael-Pournin S., LeBras S., Coumalleu
              F., Babinet C., Baldacci, P.
              Unite de Biologie du Developpement, CNRS URA 1960, Institut
              Pasteur, 25 rue du Dr Roux, 75724 Paris cedex 15, France.
              Location/Qualifiers
              1. 240425
              /organism="Mus musculus"
              /mol_type="genomic DNA"

```





# REFERENCE AUTHORS TITLE JOURNAL

3 (bases 1 to 259772)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GONR  
Center clone name: CH230-36G6  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 233717 bases at least Q40  
Consensus quality: 236719 bases at least Q30  
Consensus quality: 238622 bases at least Q20  
Estimated insert size: 252037; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces  
is believed to be correct as given, however the sizes  
of the gaps between them are based on estimates that have  
been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
the accession number will be preserved.  
\* 1 259772: contig of 259772 bp in length.

## FEATURES

### source

1. .259772  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-36G6"

### misc\_feature

1. .1403  
/note="wgs contig"  
3446. .6094  
/note="wgs contig"  
254751. .255855  
/note="wgs contig"  
258597. .259772  
/note="wgs contig"

## ORIGIN

Query Match 7.9%; Score 47.4; DB 2; Length 259772;  
Best Local Similarity 51.2%; Pred. No. 0.071;  
Matches 111; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
QY 262 TGCATGTTTAGAGATTCAGAGATAGTTTCAGGAATGTAATTTATGATTTTATACAT 321  
Db 120070 TGCATGTTTAGAGATTAACATCATGATTCACGACAAATGACAGACATGTCGTAAT 120011  
QY 322 ACACAGAGAAATACGCGACTAGTGAGAAGCTATTGCGCATGTCCTCCAGCAGAGATGATGAA 381

Db 120010 ACATAGAGCAATACATACACAAGAGACTGGAGTGAAGGGTGGCGGACCATACAGGAC 119951  
QY 382 GGCCTAAATATGAGGACCAAGAGGCGAGCAATGAAGTAATGAGCCATGAGGAGTGAATGCT 441  
Db 119950 AGTTTATACAAATTTGTTCAGCCAGACAGGAGAAAGGATACACCGTTCCCAAGGTCAACACT 119891  
QY 442 GCATGTTGTAATGAGGAGGAGAAAGACCTGTGACTTCA 478  
Db 119890 CCATGCTACAGCTACATCGAGAACCTGTAAATTTCA 119854

## RESULT 15 AC128200/c

### LOCUS DEFINITION

AC128200/c

AC128200

HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 239429)

Muzny, D. Marie, Metzker, M. Lee, Abramzon, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhr, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Greggorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guervara, W.,

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,

Hollins, B., Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, B., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensu, L., Louie, H., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,

Nwaekeme, O., Okunolu, G., Olarnpungsoo, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,

Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D.,

Shed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Tamani, K.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Tjebk, Z., Umani, K.,

Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 397.652 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-1

Perfect score: 600

Sequence: 1 taggaaattataggatcat.....cctggaatcatcattttaa 600

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001s:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	600	100.0	600	ADH13938	Adh13938 Human mon
2	600	100.0	11793	AD003803	Ado03803 Human Ccl
3	41.2	6.9	78268	ACN44342	Acn44342 Human gen
C 4	40.8	6.8	2000	ADA71938	Ada71938 Rice gene
C 5	39.4	6.6	126974	ADN30166	Adn30166 Hepatocyt
C 6	39.4	6.6	144035	ACN44062	Acn44062 Human gen
C 7	38.8	6.5	1768	ADR64667	Adr64667 Cotton cd
8	38.8	6.5	7340	AAD28378	Aad28378 Human che
9	38.4	6.4	8943	ABK39967	Abk39967 Human che
10	38.2	6.4	52302	ADA02738	Ada02738 Human CCN
11	38.2	6.4	52302	ADB72476	Adb72476 Human CCN
12	38.2	6.4	52302	ADC85218	Adc85218 Human Ccn
13	38.2	6.4	52302	ADM74333	Adm74333 Human car
14	37.4	6.2	2000	ADA71938	Ada71938 Rice gene
15	37	6.2	2646	ADR07244	Adr07244 Full leng
16	37	6.2	7261	AAS46669	Aas46669 Tumour su
17	37	6.2	101616	ACN44934_3	Continuation (4 of
18	37	6.2	110000	ACN44934_2	Continuation (3 of
C 19	36.6	6.1	201143	ABK83568	Abk83568 Human DNA
20	36.4	6.1	3364	ABL45715	Ab145715 Human mat

21	36.4	6.1	201766	10	ADL13771	Adl13771 Osteoearth
C 22	36.2	6.0	32191	4	AAS30497	Aas30497 DNA encod
C 23	36.2	6.0	32191	4	AAL06277	Aal06277 Human rep
C 24	36.2	6.0	46107	4	AAK71730	Aak71730 Human imm
25	36	6.0	1767	10	ADB63552	Adb63552 Human cdn
26	35.8	6.0	421	2	AAV88382	Aav88382 EST clone
C 27	35.8	6.0	2000	10	ADC08421	Adc08421 Rice DNA
C 28	35.8	6.0	2000	12	ADJ41665	Adj41665 Plant cdn
C 29	35.4	5.9	41787	13	ABD33599	Abd33599 Human can
C 30	35.2	5.9	1245	9	ADA31548	Ada31548 DNA encod
C 31	35.2	5.9	9729	6	ABQ76227	Abq76227 Human SCC
C 32	35.2	5.9	9729	9	ACC72422	Acc72422 Human ova
C 33	35.2	5.9	9909	9	ACC72421	Acc72421 Human ova
C 34	35.2	5.9	9909	13	ADR72882	Adr72882 Human ova
C 35	35.2	5.9	44585	10	ADC86636	Adc86636 Human GPC
36	35.2	5.9	110000	6	ABQ74964_5	Continuation (6 of
C 37	35.2	5.9	209484	11	ACN44126	Acn44126 Human gen
38	35	5.8	8495	6	ABT08488	Abt08488 Human nov
39	35	5.8	8495	12	ADH71331	Adh71331 Human gen
40	35	5.8	8495	12	ADO09837	Ado09837 Human NOV
41	34.8	5.8	960	8	ACA30251	Aca30251 Prokaryot
42	34.8	5.8	2630	5	AAS69830	Aas69830 DNA encod
43	34.8	5.8	3558	5	AAS72863	Aas72863 DNA encod
44	34.8	5.8	5807	6	ABL33155	Ab133155 Human imm
C 45	34.6	5.8	871	13	ADR60929	Adr60929 Cotton cd

## ALIGNMENTS

## RESULT 1

ADH13938

ID ADH13938 standard; DNA; 600 BP.

XX AC ADH13938;

XX DT 11-MAR-2004 (first entry)

XX DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:1.

XX KW ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytotstatic; antiarteriosclerotic; gene therapy; atherosclerosis; cancer;

XX KW cervical carcinoma.

XX OS Homo sapiens.

XX PN BP1170372-Al.

XX PD 09-JAN-2002.

XX PF 06-JUL-2000; 2000EP-00114560.

XX PR 06-JUL-2000; 2000EP-00114560.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Roel F, Soto U, Coy J, Finzer P, Delius H, Poustka A;

XX Zur Hausen H, Patzelt A;

XX WPI; 2002-165895/22.

XX PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1, useful in gene therapy, for treating atherosclerosis and cancer.

XX PS Disclosure; SEQ ID NO 1; 30pp; English.

XX The invention relates to a novel nucleic acid molecule (I) comprising a sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein having the biological activity of MCP-1. A protein encoded by a nucleic acid of the invention has cytostatic, and antiarteriosclerotic activity.

XX CC A nucleic acid of the invention may have a use in gene therapy. A compound of the invention is useful in the preparation of a medicament for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

CC	pharmaceutical composition of the invention is useful for the treatment
CC	of diseases associated with dysregulation of MCP-1 expression, e.g.
CC	atherosclerosis or cancer. The present sequence is used in the
CC	exemplification of the invention.
XX	
SQ	Sequence 600 BP; 213 A; 101 C; 150 G; 136 T; 0 U; 0 Other;
	Query Match            100.0%; Score 600; DB 6; Length 600;
	Best Local Similarity 100.0%; Pred. No. 2.3e-173;
	Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TAGGAAAATTATAGGATCATTTAGAAGCAGAAAGAGGTGGGAGCAATACCTCGAGG 60
Dd	1 TAGGAAAATTATAGGATCATTTAGAAGCAGAAAGAGGTGGGAGCAATACCTCGAGG 60
QY	61 TAGAANAATGGTGATGTGFACATCAAGCAGGGAGAAAAACCAATGAACACGATCGGAATT 120
Dd	61 TAGAANAATGGTGATGTGFACATCAAGCAGGGAGAAAAACCAATGAACACGATCGGAATT 120
QY	121 CGGGCCCCACCAAATGCATAGGAGTACAATTAAGAAGGAAGGTTTCAGTCAAGGGATTG 180
Dd	121 CGGGCCCCACCAAATGTCAGGGATGCAATTTAGAAGGAAGGTTTCAGTCAAGGGATTG 180
QY	181 AATGTTAGGGTGAAAAGTTACTCTCAACTCTCTAGTGTAAAAAGGAAACGTTTGAATACT 240
Dd	181 AATGTTAGGGTGAAAAGTTACTCTCAACTCTCTAGTGTAAAAAGGAAACGTTTGAATACT 240
QY	241 TCAGTCCAATCAGGAGGGATGTGCCATGTTTAGAGATTTCAGAGATAAGTTTCAGGAAATG 300
Dd	241 TCAGTCCAATCAGGAGGGATGTGCCATGTTTAGAGATTTCAGAGATAAGTTTCAGGAAATG 300
QY	301 TAACTTATAGATTTTTATACATACACAGAGAAATACGCACTAGTGAAGAGCTATTGCCCATG 360
Dd	301 TAACTTATAGATTTTTATACATACACAGAGAAATACGCACTAGTGAAGAGCTATTGCCCATG 360
QY	361 GTCCAAGCAAGAGATGATGAAGGGCTAAATATATGAGCCAAAGAGGCAGCAATGAAGAATG 420
Dd	361 GTCCAAGCAAGAGATGATGAAGGGCTAAATATATGAGCCAAAGAGGCAGCAATGAAGAATG 420
QY	421 AGCCATGCAAGGTTGAATGCTGCATGTTGTAATGGAGGAGAAAGACCCTGCACTTCAGA 480
Dd	421 AGCCATGCAAGGTTGAATGCTGCATGTTGTAATGGAGGAGAAAGACCCTGCACTTCAGA 480
QY	481 TATGAAAACCTCATCTTCAACCCACATTTTAAGGGGGCAGTTCCCTGAAACCGAATGCT 540
Dd	481 TATGAAAACCTCATCTTCAACCCACATTTTAAGGGGGCAGTTCCCTGAAACCGAATGCT 540
QY	541 GTTTCCTTCGATTAATACCCCATCCCAATCTCAGGCACCTCGGAATCATCCATTTAAA 600
Dd	541 GTTTCCTTCGATTAATACCCCATCCCAATCTCAGGCACCTCGGAATCATCCATTTAAA 600
	RESULT 2
ID	AD003803 standard; DNA; 11793 BP.
XX	
AC	ADO03803;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human Cc12 gene and enhancer region DNA Seqid 4.
XX	human; ds; animal model; age-related macular degeneration; AMD;
KW	gene knockout; Ccl12-deficient; Ccr2-deficient; drusen;
KW	lipofuscin accumulation; Bruch's membrane; retinal degeneration;
KW	choroidal neovascularisation; ophthalmological; gene therapy.
XX	
OS	Homo sapiens.
XX	
FN	WO2004041160-A2.
XX	
PX	21-MAY-2004.
Dd	

PF	16-OCT-2003; 2003WO-US032933.
XX	
PR	30-OCT-2002; 2002US-0422096P.
XX	
PA	(KENT ) UNIV KENTUCKY RES FOUND.
XX	
PI	Ambati J;
XX	
DR	WPI; 2004-400512/37.
XX	
XX	Testing candidate drug for treating age-related macular degeneration, by
XX	administering drug to Ccl12-deficient, Ccl12-deficient knockout mouse, and
PT	analyzing development or regression of drusen and/or lipofuscin
PT	accumulation in eye.
XX	
PS	Disclosure; SEQ ID NO 4; 64pp; English.
XX	
XX	This invention relates to a novel methods and animal models for testing
CC	candidate drugs that can be used for the treatment or prevention of age-
CC	related macular degeneration (AMD). Specifically, it refers to
CC	administering a candidate drug to gene knockout mice, in particular Ccl12-
CC	deficient, Cor2-deficient and/ or a Ccl12-deficient/Cor2-deficient dual
CC	knockout mouse. The present invention describes analysing the knockout
CC	mouse eye for development or regression of drusen and/ or lipofuscin
CC	accumulation, as well as for the effect of a candidate drug on Bruch's
CC	membrane, retinal degeneration and/ or choroidal neovascularisation.
CC	Accordingly, such compositions exhibit ophthalmological activities and
CC	can be used for gene therapy purposes. This polynucleotide sequence is
CC	the human Ccl12 gene and enhancer region DNA of the invention.
XX	
SQ	Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;
	Query Match
	Best Local Similarity 100.0%; Score 600; DB 12; Length 11793;
	Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TAGGAAATTTAGCATCATTAAGAAAGAGAGAGAGTGGAGCAAAATACCTGGAGG 60
DB	9979 TAGGAAATTTAGGATCATTTAAGAAAGAGAGAGAGTGGAGCAAAATACCTGGAGG 10038
QY	61 TAGAAATGGTGATGATGTGTACATCAAGCAGGAGAAAACCAATGAACAGATGCGAATT 120
DB	10039 TAGAAATGGTGATGATGTGTACATCAAGCAGGAGAAAACCAATGAACAGATGCGAATT 10098
QY	121 CGGGCCCAACCAATGTCAAGGATGCAAAATTAGAAAGAGAGTGTGATCAAGGGAATTTG 180
DB	10099 CGGGCCCAACCAAAATGTCAGGAGTGCACAAATTAGAAAGAGAGTGTGATCAAGGGAATTTG 10158
QY	181 AATGTTAGGGTGAAAGTTTACTACTCAACTCTGTAGTTTAAAGGAAACGTTGAGATCT 240
DB	10159 AATGTTAGGGTGAAAGTTTACTACTCAACTCTGTAGTTTAAAGGAAACGTTGAGATCT 10218
QY	241 TCAGTCCAATGAGGAGGATGTGCCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG 300
DB	10219 TCAGTCCAATGAGCAGGATGTGCCATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATG 10278
QY	301 TAATTTTAGATTTTATACATACACAGAAATACCGACTAGTGTAGAAGCTATTGCCATG 360
DB	10279 TAATTTTAGATTTTATACATACACAGAAATACCGACTAGTGTAGAAGCTATTGCCATG 10338
QY	361 GTCCAAAGCAAGAGATGATGAAGCCCTAAATATATCGCCAAAGGAGCGCAATCAAGAATG 420
DB	10339 GTCCAAAGCAAGAGATGATGAAGCCCTAAATATATCGCCAAAGGAGCGCAATCAAGAATG 10398
QY	421 AGCCATCGAGGGTGAAATGCTGCATGTTGTAATTGGAGGAGAAAGACCTGTGACTTCAGA 480
DB	10399 AGCCATCGAGGGTGAAATGCTGCATGTTGTAATTGGAGGAGAAAGACCTGTGACTTCAGA 10458
QY	481 TATGAAACCTCATCTTTCAACCCCAATTTTAAAGGGGCGAGCTTCCCTGAAACCGAATGT 540
DB	10459 TATGAAACCTCATCTTTCAACCCCAATTTTAAAGGGGCGAGCTTCCCTGAAACCGAATGT 10518
QY	541 GTTTTCCCTCCATTTACTATATACCCCATCTCAGGCACCTCGGAATCATCCATTTAA 600

Db 10519 GTTTCCTCATTACTATATACCCCAATCTCAGGCACCTGGAATCATCATTTAAA 10578  
|||||

RESULT 3  
ACN44342  
ID ACN44342 standard; DNA; 78268 BP.  
XX ACN44342;  
XX  
XX 18-NOV-2004 (first entry)  
XX Human genomic sequence hCG17932.  
XX  
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO2003073826-A2.  
XX  
XX 12-SEP-2003.  
XX  
XX 28-FEB-2003; 2003WO-US006235.  
XX  
XX 01-MAR-2002; 2002US-00087192.  
XX (SAGR-) SAGRES DISCOVERY.  
XX  
XX Morris DW;  
XX  
XX WPI; 2003-328604/31.  
XX  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
XX comprises a nucleotide sequence.  
XX  
XX Claim 1; SEQ ID NO 742; Opp; English.  
XX  
XX The present invention relates to novel DNA and protein sequences which  
XX are associated with carcinomas. The sequences are useful for: (i) for  
XX screening drug candidates; (ii) for screening of bioactive agent capable  
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
XX a bioactive agent capable of modulating the activity of CAP; (iv) for  
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
XX determining Carcinoma Associated (CA) gene copy number. In addition, the  
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of  
XX carcinoma including lymphoma. The present sequence is one such CA coding  
XX sequence. Note: This patent is an equivalent to basic patent  
XX US2002182586A1, for which no sequence data was published  
XX  
XX Sequence 78268 BP; 20402 A; 18203 C; 18818 G; 20845 T; 0 U; 0 Other;  
XX  
Query Match 6.9%; Score 41.2; DB 11; Length 78268;  
Best Local Similarity 62.7%; Pred. No. 0.55;  
Matches 64; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
XX  
QY 26 AAGCAGAAGGAGAGTGGGAGCAATACTCGAGGTAGAAATGGTGATGTGTACATC 85  
Db 9409 AGGCACAAGGAAACAGCAGGTACAAAGGCTCGGAGGTCGGAATGAACCTTAATGTGTTTCA 9468  
QY 86 AAGCAGGGAGAAACCAATGAACCAAGATGCGAATTCGGGCCC 127  
Db 9469 GAACAGAAGAAACCAGTGTACCTGGAGAGCAGCCAGGGCC 9510  
XX  
RESULT 4  
ADA71938/c  
ID ADA71938 standard; DNA; 2000 BP.  
XX  
XX ADA71938;  
XX

Tue Aug 9 17:01:24 2005

Db 626 MTGMYGMYSSYKSMWTSKMSYMGKMTCTWYTSKMGSTRSKGWRSGMSRMTNRWKK 567  
QY 481 TATGAAACCTCATCTTCAACCCACATTTTAAAGGGGCGAGCTTCCCTGAAACCCAGAAATGT 540  
Db 566 MRKRYMYMKWCTWRRCMCYTRWGYTWYTTSRSEMYTGRYKARYTSKRMYMYKYRKY 507  
QY 541 GTTTCCTCCTACTTACTATACCCCTCCCAATCTCAGGCACCTGGAATCATCCATTAAA 600  
Db 506 CWYYYGYMKCSYMMRYGYCKACKCCYACMCWKAAYSGMMYWKYKSKWRMSTKYMW 447

RESULT 5  
ID ADN30166/c  
ID ADN30166 standard; DNA; 126974 BP.  
XX AC ADN30166;  
XX DT 12-AUG-2004 (first entry)  
XX DE Hepatocyte growth factor receptor associated DNA #2.  
XX KW cytosstatic; hepatocyte growth factor receptor;  
XX KW hyperproliferative disorder; antisense technology; human; ds.  
XX OS Homo sapiens.  
XX PN US2004102622-A1.  
XX PD 27-MAY-2004.  
XX PF 23-NOV-2002; 2002US-00304019.  
XX PR 23-NOV-2002; 2002US-00304019.  
XX PA (ISIS-) ISIS PHARM INC.  
XX PI Dean NM, Bennett CF, Dobie KW;  
XX WPI; 2004-399741/37.  
XX New compound targeted to a nucleic acid molecule encoding hepatocyte  
PT growth factor receptor, useful in diagnosing and treating  
PT hyperproliferative disorder.  
PS Example 15; SEQ ID NO 12; 116pp; English.  
XX The invention describes a new compound 8-80 nucleobases in length  
CC targeted to a nucleic acid molecule encoding hepatocyte growth factor  
CC receptor, where the compound specifically hybridizes with the nucleic  
CC acid molecule encoding hepatocyte growth factor receptor comprising a  
CC sequence of 4586 bp (SEQ ID NO: 4) and inhibits the expression of  
CC hepatocyte growth factor receptor. Also described are: a method of  
CC inhibiting the expression of hepatocyte growth factor receptor in cells  
CC or tissues; screening for a modulator of hepatocyte growth factor  
CC receptor; a diagnostic method for identifying a disease state; a kit or  
CC assay device comprising the compound; and treating an animal having a  
CC disease or condition associated with hepatocyte growth factor receptor.  
CC The compound and methods are useful in diagnosing and treating  
CC hyperproliferative disorder. This sequence represents a human hepatocyte  
CC growth factor receptor associated polynucleotide.  
XX Sequence 126974 BP; 38160 A; 25346 C; 24426 G; 39042 T; 0 U; 0 Other;

Query Match 6.6%; Score 39.4; DB 12; Length 126974;  
Best Local Similarity 50.8%; Pred. No. 2.5;  
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 287 AGTTTCAGGAATCTAATCTATAGATTTTATACATACACAGAAATACGGACTAGTGAG 346  
Db 88257 AATTACACAAATTAATATTTTACCTCTGAAAAAATACAAAGAAACCGAACTTAATA 88198  
QY 347 AAGCTATTGCCATGGTCCAAGCAAGAGATGATGAAGCCCTTAATATGAGCCAAAGAGGC 406

Db 88197 TTGTTGCTAGACTGATAGAAATTAACGAAACAAGCCAGATTGAACCTGGTCCATGAGAT 88138  
QY 407 ASCAATGAAGATGAGCCATGAGGGTGAATGCTGCAATGTTGTAATGAGGAGAAAGA 466  
Db 88137 ATCAATTCAAGAGAGTCTGGCATAAAGAAAGCTGTGTGTAAGACGAGCAGGACAAAA 88078  
QY 467 CCTGT 471  
Db 88077 CATAT 88073

RESULT 6  
ID ACN44062/c  
ID ACN44062 standard; DNA; 144035 BP.  
XX AC ACN44062;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human genomic sequence HCG38705.  
XX KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.  
XX OS Homo sapiens.  
XX PN WO2003073826-A2.  
XX PD 12-SEP-2003.  
XX PF 28-FEB-2003; 2003WO-US006235.  
XX PR 01-MAR-2002; 2002US-00087192.  
XX PA (SAGR-) SAGRES DISCOVERY.  
XX PI Morris DW;  
XX WPI; 2003-328604/31.  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
PS Claim 1; SEQ ID NO 322; Opp; English.

The present invention relates to novel DNA and protein sequences which  
are associated with carcinomas. The sequences are useful for: (i) for  
screening drug candidates; (ii) for screening of bioactive agent capable  
of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
a bioactive agent capable of modulating the activity of CAP; (iv) for  
evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
carcinoma; (vi) for neutralizing the effect of CAP; (vii) for treating  
carcinoma; (viii) for inhibiting the effect of CAP; (ix) as a biochip;  
(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
determining Carcinoma Associated (CA) gene copy number. In addition, the  
CA genes are useful as DNA vaccines and the CAP are useful as markers of  
carcinoma including lymphoma. The present sequence is one such CA coding  
sequence. Note: This patent is an equivalent to basic patent  
US2002182586A1, for which no sequence data was published  
XX Sequence 144035 BP; 43371 A; 28888 C; 27578 G; 44198 T; 0 U; 0 Other;  
Query Match 6.6%; Score 39.4; DB 11; Length 144035;  
Best Local Similarity 50.8%; Pred. No. 2.6;  
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 287 AGTTTCAGGAATCTAATCTATAGATTTTATACATACAGAAATACGGACTAGTGAG 346  
Db 96760 AATTACACAAATTAATATTTTACCTCTGAAAAAATACAAAGAAACCGAACTTAATA 96701  
QY 347 AAGCTATTGCCATGGTCCAAGCAAGAGATGATGAAGCCCTTAATATGAGCCAAAGAGGC 406  
Db 96700 TTGTTGCTAGACTGATAGATTAACGAAACAAGGCAGATTGAACCTGGTCCAATGAGAT 96641



Tue Aug 9 17:01:24 2005

pharmacogenomics and for therapy of diseases e.g. cancer.

PT Novel nucleic acid useful for diagnosis and therapy of behavioral  
PT disorder, neurological disorder and cancer, comprises a sequence of a  
PT segment of chemically pretreated DNA of adrenergic alpha-1C-receptor  
PT gene.  
XX  
XX Claim 1; Page 94-98; 190pp; English.  
PS  
XX The invention relates to nucleic acids comprising a segment of chemically  
XX pretreated DNA of adrenergic alpha-1C-receptor gene. The invention also  
XX relates to oligonucleotides or peptide nucleic acid (PNA) oligomers  
XX useful for detecting cytosine methylations. The pretreated DNA is useful  
XX for the diagnosis or therapy of behavioural disorders, neurological  
XX disorders and cancer, in particular major depressive disorder, Tourette's  
XX syndrome, schizophrenia, psychiatric and neurological disorders, smoking,  
XX drug abuse, alcoholism, personality traits, compulsive gambling, human  
XX immunodeficiency virus dementia, migraine, behaviour in patients with  
XX and schizophrenia. The nucleic acid is useful for detecting the methylation  
XX state of all CpG dinucleotides and/or single nucleotide polymorphisms  
XX (SNPs). The present sequence is human chemically treated genomic DNA  
XX (SNPs).  
SQ Sequence 7340 BP; 1725 A; 249 C; 1982 G; 3384 T; 0 U; 0 Other;  
Query Match 6.5%; Score 38.8; DB 6; Length 7340;  
Best Local Similarity 54.1%; Pred. No. 1;  
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
QY 158 GGAAGTTGAGTCAAGGGAATTTGAATGTTAGGGTGAAGTACTCACTCTGAGG 217  
Db 1182 GGTAGTTTGTAGTGGAAATTTAGATGTTAGGGTGAAGTACTCACTCTGAGG 217  
QY 218 TTAAGAAGAAAGTTGAGAACTTTCAGTCCCAATCAGGAGGAGTGCCATGTTTAGAGAT 277  
Db 1242 TTTATATGATAAGTTTCAGAAATAGAAATATATTTTAAATTTGTTTATTAGTAAAGGG 1301  
QY 278 TCAGAGATAAGTTTCAGGAAATGTAA 303  
Db 1302 AAGAAGAAGAAATGAATAAAGTTAA 1327  
RESULT 9  
ABK39967  
ID ABK39967 standard; DNA; 8943 BP.  
XX  
XX ABK39967;  
XX  
XX 21-MAY-2002 (first entry)  
XX Human chemically pretreated gene sequence #24 strand 2.  
XX  
XX Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;  
KW Cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;  
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.  
XX  
XX Homo sapiens.  
XX  
XX WO200202806-A2.  
XX  
XX 10-JAN-2002.  
XX  
XX 29-JUN-2001; 2001WO-EP007470.  
XX  
XX 30-JUN-2000; 2000DE-01032529.  
XX  
XX 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPiG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-154757/20.  
XX  
XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,  
PT useful for detecting cytosine methylation state of genes associated with

PT  
XX  
PS Claim 1; SEQ ID NO 48; 24pp; English.  
XX  
XX The invention relates to a nucleic acid comprising a sequence at least 18  
XX bases in length of a segment of the chemically pretreated DNA of genes  
XX associated with pharmacogenomics according to one of the sequences of the  
XX genes ALDH6 (NM 000693), CYP11A (NM 000781), CYP11B1 (NM 000497), CYP3A3  
XX (NM 000776 and NM 017460), DPYD (NM 000110), EPHX2 (NM 001979), OCLN  
XX (NM 002538), TXNRD1 (NM 003330), UGT8 (NM 003360), MRP (NM 004996),  
XX NM 019900, NM 019901, NM 019902, NM 019862, NM 019898, NM 019899 and  
XX their complementary sequences, or a sequence (S1) chosen from 87  
XX sequences and their complements. The chemical pretreatment is bisulphite  
XX treatment to convert cytosines (but not methyl-cytosines) into uracil.  
XX Also included are an oligomer (PNA)-oligomer, comprising in each case at least one  
XX base sequence having a length of 9 nucleotides which hybridises to or is  
XX identical to a chemically pretreated DNA of genes associated with  
XX pharmacogenomics and their complements, arranged in an array for  
XX analysing diseases associated with the methylation state (CpG) and/or  
XX detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The  
XX oligomers may also be used as PCR primers. The set of 87 nucleic acids  
XX and their complements is useful for diagnosis and therapy of solid  
XX tumours and cancer. The present sequence represents one the 87 DNA  
XX sequences or its complement. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 8943 BP; 2415 A; 42 C; 1943 G; 4543 T; 0 U; 0 Other;  
Query Match 6.4%; Score 38.4; DB 6; Length 8943;  
Best Local Similarity 52.5%; Pred. No. 1.5;  
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;  
QY 183 TGTTAGGTTGAAAGTTTACTCACTCTGAGTGTAAAGGAAACGTTGAGAACTTC 242  
Db 3198 TATTGGTGAGAGATTTATTTTTTTTTTTTGTAAATTTAAAGGAAATGTAATTTAA 3257  
QY 243 AGTCCAATGAGAGGGATGTCCTCATGTTAGAGATTCAGAGATAAGTTTCAGGAAATGTA 302  
Db 3258 ATTTTGGTTTGGTAAAGTTGGGTATATAAATTATGTTTATTAGATTATTTTAAATGTT 3317  
QY 303 ACTTATAGATTTTATACATACACAGAGAAATACGAGACTAG 342  
Db 3318 GTTTTAAATTTTTTTTTTTTTTTTATTAAAGATGAATGGGGAG 3357  
RESULT 10  
ADA02738  
ID ADA02738 standard; DNA; 52302 BP.  
XX  
XX ADA02738;  
XX  
XX 06-NOV-2003 (first entry)  
XX  
XX Human CCND2 carcinoma associated gene, SEQ ID NO:1256.  
XX  
XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
KW gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO2003057146-A2.  
XX  
XX 17-JUL-2003.  
XX  
XX 26-DEC-2002; 2002WO-US041414.  
XX  
XX 26-DEC-2001; 2001US-00035832.  
XX  
XX (SAGR-) SAGRES DISCOVERY.  
XX

XX Morris DW;  
 PI  
 XX  
 DR WPI; 2003-587068/55.  
 XX  
 XX New recombinant nucleic acid encoding carcinoma associated protein,  
 PT useful for preparing compositions for treating carcinomas.  
 XX  
 XX Claim 1; SEQ ID NO 1256; 245pp; English.  
 XX  
 XX The invention relates to recombinant carcinoma associated (CA) nucleic  
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to  
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The  
 CC invention also encompasses expression vectors and host cells comprising a  
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically  
 CC binds to the protein, and a biochip comprising CA nucleic acid or  
 CC fragments thereof. The sequences of the invention were identified using  
 CC oncogenic retroviruses, which insert into the genome of the host organism  
 CC at random. Many of these do not carry transduced host oncogenes or  
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a  
 CC direct consequence of the effects of proviral integration into host  
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose  
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or  
 CC leukaemia) or a propensity to carcinoma by determination of the sequence  
 CC of a CA gene, or by determination of CA gene expression in particular  
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as  
 CC therapeutic agents and in screening and evaluating drug candidates. The  
 CC present sequence represents a specifically claimed human CA nucleic acid  
 CC sequence of the invention. Note: The complete sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 52302 BP; 13145 A; 12501 C; 12728 G; 13888 T; 0 U; 40 Other;  
 SQ

Query Match 6.4%; Score 38.2; DB 9; Length 52302;  
 Best Local Similarity 69.3%; Pred. No. 3.8;  
 Matches 52; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 27 AGCAGAGGAGGAGTGGGAGCAATACCTGGAGGTAGAATGGTGATGTGATCATCA 86  
 DB 50950 AGCAGAGGAGGAGTGGGAGCAATACCTGGAGGTAGAATGGTGATGTGATCATCA 86  
 QY 87 AGCAGGGGAGGAGCAAC 101  
 DB 51010 AACAGAGAGAGGCC 51024

RESULT 11  
 ID ADB72476  
 XX ADB72476 standard; DNA; 52302 BP.  
 AC ADB72476;  
 XX  
 XX 04-DEC-2003 (first entry)  
 DT Human CCND2 gene.  
 XX  
 XX human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;  
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003008583-A2.  
 PN  
 XX 30-JAN-2003.  
 PD  
 XX 26-DEC-2001; 2001WO-US051291.  
 XX  
 XX 02-MAR-2001; 2001US-00798586.  
 PR 23-OCT-2001; 2001US-00004113.  
 PR 08-NOV-2001; 2001US-00052482.  
 PR 30-NOV-2001; 2001US-00997722.

PR 20-DEC-2001; 2001US-00034650.  
 XX  
 XX (SAGR-) SAGRES DISCOVERY.  
 XX  
 PI Morris DW, Engelhard EK;  
 XX  
 XX WPI; 2003-239337/23.  
 XX  
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,  
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.  
 PT  
 XX Claim 1; SEQ ID NO 304; 2304pp; English.  
 XX  
 XX The invention relates to a novel recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the 660 sequences fully defined  
 CC in the specification. A polynucleotide of the invention has cytostatic  
 CC activity, and may have a use in gene therapy, or in a vaccine. The  
 CC recombinant nucleic acids and polypeptides are useful for treating  
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and  
 CC sarcomas. The present sequence represents a human gene of the invention.  
 XX  
 XX Sequence 52302 BP; 13145 A; 12501 C; 12728 G; 13888 T; 0 U; 40 Other;  
 SQ

Query Match 6.4%; Score 38.2; DB 10; Length 52302;  
 Best Local Similarity 69.3%; Pred. No. 3.8;  
 Matches 52; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 27 AGCAGAGGAGGAGTGGGAGCAATACCTGGAGGTAGAATGGTGATGTGATCATCA 86  
 DB 50950 AGCAGAGGAGGAGTGGGAGCAATACCTGGAGGTAGAATGGTGATGTGATCATCA 86  
 QY 87 AGCAGGGGAGGAGCAAC 101  
 DB 51010 AACAGAGAGAGGCC 51024

RESULT 12  
 ID ADC85218  
 XX ADC85218 standard; DNA; 52302 BP.  
 AC ADC85218;  
 XX  
 XX 01-JAN-2004 (first entry)  
 DT Human Ccnd2 genomic sequence.  
 XX  
 XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;  
 KW secreted; transmembrane; intracellular; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003045230-A2.  
 PN  
 XX 05-JUN-2003.  
 PD  
 XX 02-DEC-2002; 2002WO-US038582.  
 XX  
 XX 30-NOV-2001; 2001US-00997722.  
 PR  
 XX (SAGR-) SAGRES DISCOVERY.  
 PA  
 XX Morris DW, Engelhard EK;  
 XX  
 XX WPI; 2003-513603/48.  
 DR  
 XX  
 XX New recombinant nucleic acid comprising a nucleotide sequence of any of  
 PT the carcinoma-associated (CA) genes, useful for screening for drug  
 PT candidates for diagnosing or treating carcinomas.  
 XX  
 XX Claim 1; SEQ ID NO 4; 983pp; English.  
 PS  
 XX The invention relates to a recombinant nucleic acid comprising a  
 CC nucleotide sequence selected from any of the fully defined carcinoma-

CC	associated (CA) genes from the 50 tables given in the specification. The
CC	CA proteins are secreted, transmembrane or intracellular proteins. The
CC	recombinant nucleic acids are useful for screening for drug candidates
CC	for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC	ADC85514 represent CA genes of the invention.
XX	
XX	Sequence 52302 BP; 13145 A; 12501 C; 12728 G; 13888 T; 0 U; 40 Other;
XX	
XX	Query Match 6.4%; Score 38.2; DB 10; Length 52302;
XX	Best Local Similarity 69.3%; Pred. No. 3.8;
XX	Matches 52; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX	
QY	27 AGGAGAAGAGAGTGGGAGCAATACTCTGGAGGTAGAAATGGTGATGATGTACATCA 86
DB	50950 AGCAGAAGGAACAGAGAATGCAAGATCGGAGGTAGGATGAGCATGACGTGGTCTACG 51099
QY	87 AGCAGGGAGAAACC 101
DB	51010 AACAGAGAGAGAGCCCC 51024
XX	
XX	RESULT 13
XX	ADM74333
XX	ID ADM74333 standard; DNA; 52302 BP.
XX	AC ADM74333;
XX	AC AC
XX	DT 01-JUL-2004 (first entry)
XX	XX Human carcinoma associated (CA) nucleic acid #1.
XX	DE Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
XX	KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
XX	KW cytostatic.
XX	XX Homo sapiens.
XX	OS
XX	XX US2004072154-A1.
XX	PN
XX	PD 15-APR-2004.
XX	XX 30-NOV-2001; 2001US-00997722.
XX	PF
XX	XX 22-DEC-2000; 2000US-00747377.
PR	PR 02-MAR-2001; 2001US-00798586.
XX	XX (MORR/) MORRIS D W.
PA	PA (ENGE/) ENGELHARD E K.
XX	XX Morris DW, Engelhard EK;
PI	PI
XX	XX WPI; 2004-328562/30.
XX	DR
XX	PT New carcinoma associated gene or protein, useful for preparing a
PT	composition for diagnosing or treating carcinoma e.g., leukemia or
PT	lymphoma.
XX	XX
XX	Claim 1; SEQ ID NO 4; 29pp; English.
PS	XX
XX	XX The invention relates to new recombinant nucleic acids. The invention
CC	also relates to a host cell comprising a recombinant nucleic acid or
CC	expression vector, an expression vector comprising a recombinant nucleic
CC	acid, a recombinant protein, a method of screening for drug candidates, a
CC	method of screening for a bioactive agent capable of binding to a
CC	carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
CC	method of screening for a bioactive agent capable of modulating the
CC	activity of a CAP, a method of evaluating the effect of a candidate
CC	carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
CC	the activity of a CAP, a method of treating carcinomas, a method of
CC	neutralising the effect of a CAP and a method of diagnosing carcinoma or
CC	preparability to carcinoma. A method of evaluating the effect of a candidate
CC	carcinoma drug comprises administering the drug to a patient, removing a
CC	cell sample from the patient and determining alterations in the
CC	cell sample from the patient and determining alterations in the

expression or activation of a gene comprising the nucleotide sequence. A method of diagnosing carcinoma comprises determining the expression of one or more genes comprising the nucleic acid sequence in a first tissue type of a first individual and comparing the expression of the gene from a second normal tissue type from the first individual or a second unaffected individual, where a difference in the expression indicates that the first individual has carcinoma. A method of inhibiting the activity of a CAP comprises binding to a patient an inhibitor of CAP. Carcinomas comprises administering to a patient an inhibitor of CAP. Neutralising the effect of a CAP comprises contacting an agent specific for the CAP. The polypeptide specifically binds to the protein encoded by the nucleic acid. It comprises an antibody that specifically binds to the protein encoded by the nucleic acid. The nucleic acids are useful for preparing a composition for diagnosing or treating carcinoma e.g., leukaemia or lymphoma. This sequence represents a human carcinoma associated (CA) nucleic acid of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

Sequence 52302 BP; 13145 A; 12501 C; 12728 G; 13888 T; 0 U; 40 Other;

Query Match 6.4%; Score 38.2; DB 12; Length 52302;  
Best Local Similarity 69.3%; Pred. No. 3.8;  
Matches 52; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 27 AGGAGAAGAAAGTGGGAGCAAACTCTGGAGGTAGAAATGGTGATGATGTGTACATCA 86  
50950 AGCAGAAGGAACAGAGATGCAAAAGATCCGGAGGTAGGAATGACATGACGTGCTTACG 51009

Db 87 AGCAGGGAGAAACC 101  
51010 AACAGAGAGAGCCC 51024

RESULT 14  
ADA71938  
ID ADA71938 standard; DNA; 2000 BP.  
AC ADA71938;  
DT 20-NOV-2003 (first entry)  
DE Rice gene, SEQ ID 5263.  
EX Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.  
XX Oryza sativa.  
OS  
PN WO2003000898-A1.  
PD 03-JAN-2003.  
XX 22-JUN-2001; 2001WO-IB001105.  
PF 22-JUN-2001; 2001WO-IB001105.  
PR (SYGN ) SYNGENTA PARTICIPATIONS AG.  
PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
DR Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX Claim 27; SEQ ID NO 5263; 899pp; English.  
PS The present invention relates to a method (M1) for identifying genes  
CC



***This Page Blank (uspto)***

Result No.	Score	Query Match	Length	DB	ID	Description
1	208.6	34.8	514	9	CE463871	tigr-gss-
C 2	65.8	11.0	997	9	CNS005PE	AL060767 Drosophil
3	57	9.5	1101	9	CNS0006J	AL062049 Drosophil
4	49.6	8.3	1101	9	CNS0106X	AL098595 Drosophil
C 5	49.2	8.2	1101	9	CNS00039G	AL063921 Drosophil
C 6	47.8	8.0	987	9	CNS000418	AL066537 Drosophil
C 7	46.6	7.8	771	9	AG287979	AG287979 Mus muscu
8	45	7.5	1101	9	CNS000D1	AL065414 Drosophil
9	44.8	7.5	1101	9	CNS0182P	AL108811 Drosophil
C 10	44.4	7.4	1181	8	BZ896565	BZ896565 SP__Ba008
C 11	43	7.2	1101	9	CNS000EH	AL069378 Drosophil
C 12	42	7.0	802	9	AG411354	AG411354 Mus muscu
13	42	7.0	893	9	CNS03PV8	AL255257 Tetraodon
14	42	7.0	1065	9	AG379935	AG379935 Mus muscu
C 15	41.6	6.9	997	9	CNS0134P	AL102403 Drosophil
C 16	41.4	6.9	938	9	CNS006TJ	AL065906 Drosophil
C 17	41.4	6.9	1287	9	AG391026	AG391026 Mus muscu
18	41.2	6.9	536	1	AI544540	AI544540 fb72a11.x
19	41.2	6.9	756	9	AG077606	AG077606 Pan trogl
20	41.2	6.9	783	7	CK029219	CK029219 AGENCOURT
C 21	41.2	6.9	791	9	AG372076	AG372076 Mus muscu
C 22	41.2	6.9	1204	9	CNS01652	AL106628 Drosophil
C 23	40.8	6.8	949	8	AZ193945	AZ193945 SP_1025 A
C 24	40.8	6.8	1055	8	AG181297	AG181297 Pan trogl

61	QY	TAGAAATGGTGATGATGCTGTACATCAAGCAGGAGAAAACCAATGAACCGATCGCAATT	120
126	Db	TAGAAATGATGATGGCCCTGTACATAAAGAGGGTGAACCAAT-----CAGATGGAGCT	181
121	QY	CGGGCCACACCAATGTGTCAAGGGA---TGACAAATGAGAAAGAGGTTGAGTCAAGGAT	177
182	Db	GAAATGCTATGTCCTGTCAAAATGATGGTGAAGTTTAGAAAGGAAAGTTGAGTAAAGGATT	241
178	QY	TTGAATGTTAGGTTGAAGAAAGTTACTCTCAACTCTGTAGGTTAAAAGGAAACGTTTGAA	237
242	Db	CTGAATGTGCAGTGTAGTAAAGTCAGTACTCACCCCTCATAGTTAAACGGAGATATTGAAA	301
238	QY	TCTTCAGTCCAAATCAGAGGAGGATGTGCATGTTTAGAGATTCAAGATAAAGTTTCAGAA	297
302	Db	TGTTTGGGCCAATGAGGAAGCACATGCCATGTGTGAAGTTTCAGAGATCAG--TCAGTAA	359
298	QY	ATGTAACTTATAGATTTTATACATACACAGAGAAATACGGACTAGTGTAGAAGCTATTGCC	357
360	Db	ATGTAACTTACTTGGCTTCATCAAGCCTAGAGAAATATGCACAAAGGAGAGCTACTGCC	419
358	QY	ATGGTCCAAAGCAAGAGATGATGAAGGCCCTAAATATGGAGCCAAAGGCGAGCAAT---	414
420	Db	ATGATCCAGGCAAGAGATGATGAGTGAAGTGAATATGGAGCCCAAGAGCAGCTTTGGAGA	479
415	QY	AGAAATGAGCCATGCAGGTTGAATGCTGCATGTTG	449
480	Db	AGAAATGAATAGGTAGATGAATGAATGCTGTGGCTTG	514

RESULT 2  
 CNS005TE/c  
 LOCUS  
 Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 DEFINITION  
 BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit  
 fly) genomic survey sequence.

ACCESSION	AL060767
VERSION	AL060767.1
KEYWORDS	GI:4943573
SOURCE	GSS.
ORGANISM	<i>Drosophila melanogaster</i> (fruit fly)
	<i>Drosophila melanogaster</i>
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; <i>Drosophila</i> .
REFERENCE	1 (bases 1 to 997)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
	Bp 191 91006 EVRY cedex - FRANCE (E-mail : <a href="mailto:secref@genoscope.cns.fr">secref@genoscope.cns.fr</a> )
	- Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> )
COMMENT	Determination of this BAC-end sequence was carried out as part of a
	Genome Project (BDGP).

collaboration with the serology group at the University of California, San Diego. The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP *Drosophila melanogaster* BAC library was prepared by Kazuhiro Otsuoka and Aaron Mammone in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RP11-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2. As can be seen, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://www.fruitfly.org/bac.html>.

FEATURES	source	Location/Qualifiers
1..997		/organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACR12K22" /clone_lib="RPCI-98" /note="end : TET3"

## ORIGIN

[illegible]

```

RESULT 3
CNS0006J      1101 bp      DNA      linear      GSS 03-JUN-1999
LOCUS         Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION    BACR01M22 of RC1-98 library from Drosophila melanogaster (fruit
              fly) , genomic survey sequence.
ACCESSION     AL062049
VERSION       AL062049.1  GI:4938511
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE        Drosophila melanogaster
ORGANISM      Drosophila melanogaster
              Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
              1 (bases 1 to 1101)
REFERENCE     Direct Submission
AUTHORS       Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr ;
              Web : www.genoscope.cns.fr)
              Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP) .
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazuoto Ooeogawa and
              Aaron Mammoss in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RC1-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain v2; cn bw sp, the same strain used for the BDGP's
              pi and EST libraries. A more detailed description of the library
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
COMMENT

```



Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BOP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source  
1. 987  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_lib="BACPAC16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

ORIGIN

Query Match 8.0%; Score 47.8; DB 9; Length 987;  
Best Local Similarity 22.5%; Pred. No. 0.045;  
Matches 91; Conservative 121; Mismatches 193; Indels 0; Gaps 0;  
QY 62 AGAAATGGTGTATGTCATCAAGCAGGAGAAACCAATGAACACGATGCGAATTC 121  
Db 978 AGAAATGGTGTATGTCATCAAGCAGGAGAAACCAATGAACACGATGCGAATTC 121  
QY 122 GGGCCCAACCAATGTCAGGATGCAATTAAGAAAGGAGGTTGATGTCAGGATTCGA 181  
Db 918 RAAR 859  
QY 182 ATGTTAGGTTGAAAGTTACTACTCACTCTGTAGTTTAAAGGAAACCTTCAGAAATCTT 241  
Db 858 RRRRRGR 799  
QY 242 CAGTCCATGAGGAGGATGTCGATGTTTAGAGATTTCAGAGATAGTTTCAGGAATGT 301  
Db 798 RARGRRGR 739  
QY 302 AACTTATAGATTTTATACATACAGAGAAATACGACTAGTCAGAGAGCTATTGCCATCG 361  
Db 738 AAAAARAAGAAAR 679  
QY 362 TCCAGCAGAGATGATGAGGCTTAATATGAGCCAAAGAGGACGACGATGAAGATGA 421  
Db 678 GCGARRRRGR 619  
QY 422 GCCATGCGAGGTTGAAATGCTGATGTTTAAATGAGGAGAAAGA 466  
Db 618 GARRGR 574

RESULT 7  
LOCUS AG287979 771 bp DNA linear GSS 02-JUN-2004  
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-061N22.TJ, genomic survey sequence.  
ACCESSION AG287979.1 GI:47860856  
VERSION AG287979  
KEYWORDS GSS.  
SOURCE Mus musculus molossinus  
ORGANISM Mus musculus molossinus  
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
TITLE BAC end Sequences of Library MSMg01  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 771)  
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

FEATURES  
source  
1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone\_lib="BACPAC16"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

ORIGIN  
Query Match 8.2%; Score 49.2; DB 9; Length 1101;  
Best Local Similarity 12.9%; Pred. No. 0.019;  
Matches 59; Conservative 233; Mismatches 163; Indels 3; Gaps 1;  
QY 63 GAAATGGTGTATGTCATCAACGAGGAGAAACCAATGAACACGATGCGAATTCG 122  
Db 1100 KARRWGGDTWDRTRKDDWDTKWTTWKDRADRRWAGDADRWAWDDGAGTWTATWWM 1041  
QY 123 GGCCCAACCAATGTCAGAGGATGCAATTAAGAAAGGAGTGTGATCAAGGATTCGAA 182  
Db 1040 WWWATWDTWDDKWWATAAATGTTWRTAWRTAWRTAWRTAWRTAWRTAWRTAWRT 981  
QY 183 TGTAGGTTGAAAGTTACTACTCACTGTAGTTTAAAGGAAACCTTCAGAAATCTTC 242  
Db 980 GRRDGRKRKDK 921  
QY 243 AGTCCAAATGAGGAGGATGTCGATGTTTAGAGATTTCAGAGATAGTTTCAGGAAATGTA 302  
Db 920 DDDGAGDKDDGKGDADDTDGTDKDDDKDKDKDKDKDKDKDKDKDKDKDKDKDKDK 864  
QY 303 ACTTATAGATTTTATACATACAGAGAAATACGACTAGTCAGAGAGCTATTGCCATGT 362  
Db 863 WNGWADADWTTWDAADDDWADDDWADDDWADDDWADDDWADDDWADDDWADDDW 804  
QY 363 CCAGCAACAGATGATGAAGGCTTAATATGAGCCAAAGAGGAGGACCAATGAAGATGAG 422  
Db 803 RDRKRRADRRADRRADRRADRRADRRADRRADRRADRRADRRADRRADRRADRR 744  
QY 423 CCATGAGGTTGAAATGCTGATGTTTAAATGAGGAGGAGAAACCTTCGATTCAGATA 482  
Db 743 RDRAGTAGRKRRTWKRWRKRRDTRWDADADDDTARDRRRGGDDGADAGKKTGRKR 684  
QY 483 TGAACACCTCATCTCAACCCACCATTTTAAGGGGCGAG 520  
Db 683 RDRATWDRDADWADADWTTTDTTDDDKDKDKDKDKDKDKDKDKDKDKDKDKDKDK 646

RESULT 6  
LOCUS CNS00418 987 bp DNA linear GSS 03-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR09C16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL066537  
VERSION AL066537.1 GI:4942778  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)  
ORGANISM Drosophila melanogaster  
REFERENCE Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
TITLE BAC end Sequences of Library MSMg01  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 987)  
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
COMMENT - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and

**COMMENT**

source 1.771

## ORIGIN

[illegible]

62 AGAAATGGTGATGTGTACATCAAGCAGGAGAAAACCAATGAACCAGATGCCGAA TTC 121

Qv 122 GGGCCACACCAATGTCAAGGGATGACAA 150

122 GGGCCACACCAATGTCAGAGGGATGACAA 130  
 QY || | | | | | | | | | |

## FEATURES

**Location/Qualifiers**

Tue Aug 9 17:01:24 2005

Tel: 520 626 3967  
Fax: 520 621 9288  
Email: <http://genome.arizona.edu>  
PCR Primers  
FORWARD: atc agc ggc cgc gat cc  
BACKWARD: gta aaa cga cgg cca gtc  
Plate: 0082 row: 1 column: 11  
Seq primer: atc agc ggc cgc gat cc  
Class: BAC ends.

Location/Qualifiers  
1. .1181  
/organism="Sorghum propinquum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:132711"  
/clone\_lib="SP\_Ba0082111"  
/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII; Paterson lab BAC library (HindIII)"

FEATURES  
source

ORIGIN

Query Match 7.5%; Score 44.8; DB 9; Length 1101;  
Best Local Similarity 26.3%; Pred. No. 0.33;  
Matches 121; Conservative 129; Mismatches 209; Indels 1; Gaps 1;  
QY 5 AAAATTATAGATCAATTAAGAAAGGAGAGAGGAGGAGTGGGACCAATACCTGGAGGTAGA 64  
DB 479 AAAAAAGAGAGAGVGTGGGAAAGGGRGAGATGAGAGGAAAGGAAAGGAAAGGAAAGG 538  
QY 65 AATGGTGATGATGTATCATCAAGCAGGAGGAGAAACCAATGAACAGATGCGAATTCGGG 124  
DB 539 AAAAAAGGARGAAGGGRGREGVAGRGDAAATAAAAAAADDTKRDWARGAARAAA 598  
QY 125 CCCACCAATGCAAGGGATGCAATTTAGAAAGGAGGTTGAGTCAAGGGATTTGAATG 184  
DB 599 AADAAAARAAKWTWAAAAAARAAAAAAGGTTKTTAAAAAAGGGGKGK 658  
QY 185 TTAGGTGAAAGTTACTACTCACTCTGTAGGTTAAAGGAAACGTTGAGATCTTCAG 244  
DB 659 KAADGGGAAADAAWRGRGRRAAARAAAAADAKRAAAAAAAGGAAADAGG 718  
QY 245 TCCAAATGAGGAGGATGCGCATCTTAGAGATTTCAGAGATAAGTTTCAGGAAATGAAC 304  
DB 719 KRKKKGKDGKGGKATKTAWAKGRKDWGT-ATAWTTWDATWDTWKAATDTDKRAAG 777  
QY 305 TTATAGATTTTATACATACACAGAGAAATAGGACTAGTGAAGAGCTATTCGCATGGTCC 364  
DB 778 RKRDARTARDGGRARTRRRAAAGGKRARAGARRAARRAARDRDWDWAAAAAAA 837  
QY 365 AAGCAAGAGATGATGAAGCCTTAATATGGAGCCAAAGAGGAGCAGCAATGAAGATGAGCC 424  
DB 838 AAAAAAATTTWRDWDWDWDTRDDDDTTAAWDDARABARRRRRRRRRRARRA 897  
QY 425 ATGCAGGGTGAATGCTGCTGTTGTAATGAGGAGAAA 464  
DB 898 ADDTDTKORWADTTDKDTTKTTTDDDDWDKAKRDRWA 937

RESULT 10  
BZ696565/c 1181 bp DNA linear GSS 02-JUL-2003  
LOCUS SP\_Ba0082111.f SP\_Ba Sorghum propinquum genomic clone  
DEFINITION SP\_Ba0082111 5', genomic survey sequence.

ACCESSION BZ696565  
VERSION BZ696565.1 GI:28390715  
KEYWORDS GSS.  
SOURCE Sorghum propinquum

ORGANISM Sorghum propinquum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 1181)  
Wing, R., Yu, Y., Kim, H.R., Collura, K., Pries, G., Currie, J.,  
Soderlund, C. and Hatfield, J.  
Sequencing of Sorghum BAC ends.  
http://genome.arizona.edu/stc/sorghum

Unpublished (2003)  
Contact: Rod Wing

JOURNAL Arizona Genomics Institute  
COMMENT University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA

RESULT 11  
CNS000EMH/c 1101 bp DNA linear GSS 04-JUN-1999  
LOCUS CNS000EMH/c  
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC;  
BACR29M06 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.

ACCESSION AL069378  
VERSION AL069378.1 GI:4949521  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)  
Genoscope.  
Direct Submission

AUTHORS Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
JOURNAL - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

## COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

Location/Qualifiers  
 1..1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR29M06"  
 /clone\_lib="RPCI-98"  
 /notes="end : T7"

## ORIGIN

Query Match 7.2%; Score 43; DB 9; Length 1101;  
 Best Local Similarity 29.3%; Pred. No. 1;  
 Matches 101; Conservative 77; Mismatches 167; Indels 0; Gaps 0;

QY 144 ATGACAATTAGAAGGAGTTGAGTCAAGGAGTTGAATGTTAGGGTGAAGTTACTA 203  
 Db 891 ATATGATDAATAMWTGTGGAGTGTATGGARRRRARARAAKGTGTATGATKAKAKA 832  
 QY 204 CTCACCTCTGAGTTAAAGGAAAGTTGAGATCTTCAGTCCATGAGGAGGATGTG 263  
 Db 831 WKKKATGTGKAGATRTDRAGAKAGAGTRGRWAGTATATKATRAARAADAGATAKAG 772  
 QY 264 CCATGTTTATAGATTCAGAGATTAAGTTTCAGGAATGTAACCTTATATATATACAT 323  
 Db 771 RATRARRAKATRTAGRKAARAKAWKAWGAGTGKAWAKATGTGWRATTTKAKWRKA 712  
 QY 324 ACAGAGAAATACGGACTAGTGAAGCTATTGCGATGGTCCAGCAAGAGATGATGAAG 383  
 Db 711 RKRWGTAKATDTGAAAAAARTAAAGAGATRTTRRRKRWKAWAWATATAAGTGWGTAG 652  
 QY 384 CCTAATATGGCCAAAGGAGGAGCAATGAAGATGAGCCATGAGGCTGAAATGCTGC 443  
 Db 651 ARAAARAGAWAAGTGTGTGDDGGKKRDKRTDGGRWGTTDTAWRTAGTGTGGWKKGTGT 592  
 QY 444 ATGTTGTAATGAGGAGAAAGACCTGTGACTTCAGATATGAAAA 488  
 Db 591 WKKTGGWGTGTGKGGKWRATATAWAGTKKKTTTATAAAAAWAW 547

## RESULT 12

AG411354/c 802 bp DNA linear GSS 03-JUN-2004  
 LOCUS Mus musculus molossinus DNA, Clone:MSMg01-271024.TJ, genomic survey sequence.

## DEFINITION

AG411354

## ACCESSION

AG411354

## VERSION

AG411354.1 GI:48054040

## KEYWORDS

## SOURCE

## ORGANISM

Mus musculus molossinus  
 Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 BAC end Sequences of Library MSMg01  
 Unpublished  
 JOURNAL  
 2 (bases 1 to 802)  
 Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
 Direct Submission  
 TITLE

## JOURNAL

Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: [hattori@gscc.riken.jp](mailto:hattori@gscc.riken.jp), URL: <http://hgpc.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

## COMMENT

Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe ([abe@rtc.riken.jp](mailto:abe@rtc.riken.jp)). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
 phone: 81-298-36-9189, fax: 81-298-36-9199  
 e-mail: [abe@rtc.riken.jp](mailto:abe@rtc.riken.jp)

## PRIMERS

## Sequencing : TJ

## LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

## FEATURES

## source

Location/Qualifiers  
 1..802  
 /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSMg01-271024.TJ"  
 /sex="male"  
 /tissue type="mixture of kidney and spleen"  
 /clone\_lib="MSMg01 Mouse Male BAC Library"

## ORIGIN

Query Match 7.0%; Score 42; DB 9; Length 802;  
 Best Local Similarity 49.5%; Pred. No. 1.8;  
 Matches 108; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 10 TATAGATCATTTAAGAAAGGAGAGAGTGGAGCAAAATACCTGGAGGTAGAATGG 69  
 Db 691 TAGAGAAAGGAGAAAATAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 632  
 QY 70 TGATGATGTGTACATCAAGCAGGGAGAAACCAATGACCATCGGATTCGGCCCCAC 129  
 Db 631 TGAAGAAGGTAATAAGGGGGGAGAAAGATAATAAAGGAGAAAGGAGGAGGAGG 572  
 QY 130 ACCAATGTCAAGGGATGCAATTAGAAAGGAGAGTGTGAGTCAAGGATTTGAATGTTAG 189  
 Db 571 AGGATGAAAGGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 512  
 QY 190 GTGAAAAGTTACTACTCAACTCTGTAGGTTAAAGGAA 227  
 Db 511 GGAGGTAATGAAAAATGAGGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474

## RESULT 13

## CNS03PY8

## LOCUS

## DEFINITION

CNS03PY8 893 bp DNA linear GSS 01-SEP-2000  
 Tetraodon nigroviridis genome survey sequence T7 end of clone 046011 of library G from Tetraodon nigroviridis, genomic survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

AL255257.1 GI:7976269  
 GSS; genome survey sequence.  
 Tetraodon nigroviridis  
 Tetraodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodontidae; Tetraodon.

## REFERENCE

## AUTHORS

1  
 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissbach, J.  
 Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
 Nat. Genet. 25 (2), 235-238 (2000)  
 TITLE  
 JOURNAL



Job time : 2839.9 secs

Search completed: August 4, 2005, 14:31:57

***This Page Blank (uspro)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 123.644 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-1

Perfect score: 600

Sequence: 1 taggaataattagatcat.....cctggaatcatcattataa 600

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/PCRU COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	70.6	11.8	7218	1	US-08-232-463-14
C 2	40	6.7	832	4	US-09-621-976-2813
C 3	39.4	6.6	128175	4	US-09-949-016-16268
C 4	39.4	6.6	177231	4	US-09-949-016-15841
C 5	36.8	6.1	474	4	US-09-621-976-18033
C 6	36.4	6.1	16741	4	US-09-949-016-16431
C 7	36.2	6.0	289	3	US-09-007-005-17
C 8	36.2	6.0	289	3	US-09-244-796-17
C 9	36.2	6.0	832	4	US-09-621-976-2813
C 10	36	6.0	209210	4	US-09-949-016-15094
C 11	35.4	5.9	174639	4	US-09-949-016-16509
C 12	35.2	5.9	1245	4	US-09-328-352-2835
C 13	35.2	5.9	10321	4	US-09-949-016-13587
C 14	35.2	5.9	786431	4	US-09-751-389-3
C 15	35	5.8	84525	4	US-09-949-016-16678
C 16	35	5.8	165651	4	US-09-949-016-13032
C 17	34.6	5.8	49440	4	US-09-949-016-14150
C 18	34.4	5.7	101951	4	US-09-949-016-15648
C 19	34.4	5.7	105055	4	US-09-949-016-14001
C 20	34.2	5.7	601	4	US-09-949-016-58280
C 21	34.2	5.7	601	4	US-09-949-016-194028
C 22	34.2	5.7	76553	4	US-09-949-016-13432
C 23	34.2	5.7	85675	4	US-09-949-016-12333
C 24	34.2	5.7	85675	4	US-09-949-016-15956
C 25	34	5.7	248	3	US-09-007-005-32
C 26	34	5.7	248	3	US-09-244-796-32
C 27	34	5.7	277	3	US-09-007-005-3

Sequence 3, Appli  
Sequence 56, Appl  
Sequence 8976, Ap  
Sequence 13440, A  
Sequence 135585,  
Sequence 146580,  
Sequence 17352, A  
Sequence 14968, A  
Sequence 15858, A  
Sequence 15546, A  
Sequence 37, Appl  
Sequence 44, Appl  
Sequence 16073, A  
Sequence 8286, Ap  
Sequence 23568, A  
Sequence 2, Appli  
Sequence 15273, A  
Sequence 68086, A

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCES/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; TOPOLOGY: linear  
; STRANDEDNESS: single  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 11.8%; Score 70.6; DB 1; Length 7218;

Tue Aug 9 17:01:24 2005

118 WRKSYRTRCWAYAWKTKESYYWCVKWKCKMMMMAMAMAYGKTMVBRACWKTRYWR 59  
485 AAAACCTCATCTTCAACCCACATTTTAAAGGGGAGGAGCTCCCTGAAACAGATGT 540  
58 WWAHAWHWWTMMMYIYWRAMKRRWWRKWRWSWMMWMAWGMTWRAARWWRW 3

RESULT 3  
US-09-949-016-16268/c  
; Sequence 16268, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16268  
; LENGTH: 128175  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16268

Query Match 6.6%; Score 39.4; DB 4; Length 128175;  
Best Local Similarity 50.8%; Pred. No. 0.48;  
Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 287 AGTTTCAGGAAATGTAACCTTATAGATTTTATACATACACAGAGAAATACGACTAGTGAG 346  
Db 88760 AATTACAAAATAAATAATTTTCACCTCTGAAAATAACAGAAACCGAACTTAATA 88701

QY 347 AAGCTATTGCCATGCTCCAGCAAGAGATGATGAGCCCTAAATATGAGCCAAAGAGGC 406  
Db 88700 TTGTTGCTAGACTGATAGATTAATTAACGGAACAGAGAGATTGAACTGGTCCATGAGAT 88641

QY 407 AGCAATGAAGATGAGCCATGAGGCTGAGGCTGAAATGCTGATGTTGTAATATGAGGAGAAAGA 466  
Db 88640 ATCAATTCAAGAGAGATGCTGCGATTAAGAAAGCTGTGTGTAAGAGCAGGAGCAAAA 88581

QY 467 CCTGT 471  
Db 88580 CATAT 88576

RESULT 4  
US-09-949-016-15841  
; Sequence 15841, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15841

Best Local Similarity 3.3%; Pred. No. 1.4e-11;  
Matches 13; Conservative 237; Mismatches 141; Indels 0; Gaps 0;

QY 5 AAAATTATAGGATCATTAAGAAGAGAGAAAGAGTGGGAGCAAACTCTGGAGGTAGA 64  
Db 1457 AAAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398

QY 65 AATGGTGATGATGTACATCAAGCAGGAGAGAAACCAATGAACAGATGCGAATTCGGG 124  
Db 1397 RRR 1338

QY 125 CCACACCAATGTTCAAGGATGACAAATWAGAAAGAGTTGAGTCAAGGATTTGAATG 184  
Db 1337 RRR 1278

QY 185 TTAGGTGAAAGTTACTACTCAACTCTGAGTTTAAAGGAAACGTTGAGAAATTCAG 244  
Db 1277 RRR 1218

QY 245 TCCAATGAGGAGGATGTCATGTTTAGAGATTCAGATTAAGTTTCAAGAAATGAAC 304  
Db 1217 RRR 1158

QY 305 TTATAGATTTTATACATACACAGAGAAATACGACTAGTCAAGAGCTATTGCCATGGTCC 364  
Db 1157 RRR 1098

QY 365 AAGCAAGAGATGATCAAGGCTCAATATGGA 395  
Db 1097 RRR 1067

RESULT 2  
US-09-621-976-2813/c  
; Sequence 2813, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 235..399  
US-09-621-976-2813

Query Match 6.7%; Score 40; DB 4; Length 832;  
Best Local Similarity 10.5%; Pred. No. 0.024;  
Matches 31; Conservative 140; Mismatches 125; Indels 0; Gaps 0;

QY 245 TCCAATGAGGAGGATGTCATGTTTAGAGATTCAGAGATAAGTTTCAAGAAATGTAAC 304  
Db 298 YCYMCWCKMYRGRCAWYTWARGMWSYAWGKWSRMSMCTRMYYKKGSTYWTMKC 239

QY 305 TTATAGATTTTATACATACACAGAGAAATACGACTAGTGAAGCTATTGCCATGGTCC 364  
Db 238 TCATWCYWKYKRWMSKTCWSGSRGGYMTSYTSRYSMYWASWYTWCMWGRWST 179

QY 365 AAGCAGAGATCATCAAGCCCTAAATATGAGCCAAAGAGCAGCAATGAAGATGAGCC 424  
Db 178 YWTMAGKKWRYATWRRAMWMAATWMMWMAWCMSSRGAAMYRTMMWGRYIW 119

QY 425 ATGACGGGTGAATGCTGATGTTTAATGAGGAGAGAAACCTGTGACTTCAGATATG 484



```

; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihne
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 6.0%; Score 36.2; DB 3; Length 289;
Best Local Similarity 4.6%; Pred. No. 0.22;
Matches 9; Conservative 89; Mismatches 99; Indels 0; Gaps 0;

QY 1 TAGGAAAATTATAGGATCATTTAGAAAGGAGAGAGAGAGAAATACCTGGAGG 60
DB 28 UURURARCARARURURARCARARURGRNRRNRNRNRNRNRNRNRNRNR 87
QY 61 TAGAAATGGTGATGATGTGTACATCAAGCAGGAGAGAAACCAATGAACGATCGAATT 120
DB 88 NNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 147
QY 121 CGGGCCCCACCAATGTCAAGGGATGACAAATTAGAAAGGAGGTTGAGTCAAGGATTG 180
DB 148 NNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 207
QY 181 AATGTTAGGTTGAAAG 197
DB 208 NNRNRNRNRRCRARG 224

RESULT 9
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.J.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 6.0%; Score 36.2; DB 4; Length 832;
Best Local Similarity 13.1%; Pred. No. 0.38;
Matches 29; Conservative 102; Mismatches 90; Indels 0; Gaps 0

QY 160 AAGTTGAGTCAAGGATTTGAATGTTAGGTTGAAAAGTTACTACTCAACTCTGTAGGTT 219
DB 143 WWRKWKAWTTTWWKTTYWAAATRYWMMCMWTRWRASWWYCWGWGKRWKSTWRSY 202
QY 220 AAAAGGAAACCTTGAGAACTTTCAAGTCAATGAGGAGGATGCGCATGTTAGAGATTC 279
DB 203 ASARSAKCCYSCSWGAMSKWYWRWRGWTAGWKAWRASCMWRRKYAGKSTSYK 262
QY 280 AGAGATAAGTTTCAGGAAATGTAATCTATAGATTTTATACATACACAGAGAAATACGAC 339
DB 263 SMWCMWTRSWKYCYTKARWTGYCYRKGMMWKKGRWTYASKYKWKWKKWCMWARMYST 322
QY 340 TAGTCAGAGAGCTATTGCGCATGTTCAAGCAAGAGATGATGA 380
DB 323 GTRASWWRWRYTMMKWKYAWARAARWAWWAWRRA 363

RESULT 10

```

US-09-949-016-15094  
; Sequence 15094, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15094  
; LENGTH: 209210  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(209210)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15094

Query Match 6.0%; Score 36; DB 4; Length 209210;  
Best Local Similarity 47.0%; Pred. No. 7.5;  
Matches 111; Conservative 0; Mismatches 125; Indels 0; Gaps 0;  
  
QY 164 TTGAGTCAAGGATTGGAATGTTAGGTCGAAGTTACTACTCAACTCTGTAGTTAAAA 223  
DB 199117 TTGAGTCCAGGAGGTCAAGCTGCAGTGAGTATGTCACCTGCACCCCACTTTG 199176  
  
QY 224 GGAACGTTGAGAACTTCAGTCCATGAGGAGGATGTCCTGTTAGAGATTTCAGAG 283  
DB 199177 AGCAATAGAGTGAGACCCCTGTCTCAACAAAAAAGTGTTCCTTTTCAGTCTCAGAT 199236  
  
QY 284 ATAAGTTTCAGGAAATGTAATCTATATAGATTTTATACATACAGAGAAATACGGACTAGT 343  
DB 199237 TTCTTTTTCAGTATATAAAAAAGATTATATATCTTCTTTTCAGTAAGATTATATATCTCT 199296  
  
QY 344 GAGAAGCTATTGCCATGTCCTCAAGCAAGAGATCATGAAGCCCTAAATATGGAGCCA 399  
DB 199297 CGTCTTCTCTATTTGAGCCAACTGTGTGATGAAGCTATCAGAAATGTGAATGCA 199352

RESULT 11  
US-09-949-016-16509/c  
; Sequence 16509, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16509  
; LENGTH: 174639  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:

; NAME/KEY: misc\_feature  
; LOCATION: (1)...(174639)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16509  
  
Query Match 5.9%; Score 35.4; DB 4; Length 174639;  
Best Local Similarity 49.2%; Pred. No. 11;  
Matches 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0;  
  
QY 25 AAAGGAGAGGAGAGAGTGGGAGCAATACCTGGAGGTAGAAAATGGTGATGATGTGTACAT 84  
DB 148859 AAAGGAGACAGGAGAGAGGGAAGAAAGCAAGGGGTAGAAAGGTTAGGGGTGTAAAG 148800  
  
QY 85 CAAGCAGGAGGAGAAACCAATGAACAGATGCGAATTCGGGGCCACACCAATGTCACAGGGA 144  
DB 148799 AAAGAAAGAAAAAGAAAAAAGAGACTAGAACTACAGACAAATCAAGCTATCAAGC 148740  
  
QY 145 TCACAATTAGAAAGGAAGGTTGAGTCAAGGGATTTGAATGTTAGGGTGAAAAATTTACTAC 204  
DB 148739 AGTATAATATATATGTAACTGGAATGGGAGCAGTGCAATTTGAAGAAATAATGACATTTT 148680  
  
QY 205 TCAACTCTG 213  
DB 148679 CGAAATTTG 148671

RESULT 12  
US-09-328-352-2835/c  
; Sequence 2835, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2835  
; LENGTH: 1245  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2835

Query Match 5.9%; Score 35.2; DB 4; Length 1245;  
Best Local Similarity 50.0%; Pred. No. 0.97; Mismatches 0; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
  
QY 290 TTCAGGAAATGTAACCTATATAGATTTTATACATACACAGAGAAATACGGACTAGTGAGAAG 349  
DB 381 TGCAATGTAATAACATCGAAACATCTGAGCAAAATAAGAAAAAGCTCATTGAAATAATTTG 322  
  
QY 350 CTATTGCCATGTCCAAAGCAAGAGATGATGAAGGGCTAAATATGGAGCCAAAGAGCAGC 409  
DB 321 CAAAGAAATAGCTGCAATAATGACTGGGGCGGTCCAATATATAATCAAGAAAAAGCAAT 262  
  
QY 410 AATGAGAAATGAGCCATGCAGGTGAAATGCTGCATGTTGTTAAATGGAGAGAAAG 465  
DB 261 AATGAGTAATGAGCAAGTAGCGTAAAGCATAGGTGGCAAAAATCAAAGTCAATG 206

RESULT 13  
US-09-949-016-13587/c  
; Sequence 13587, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755

Tue Aug 9 17:01:24 2005

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13587
; LENGTH: 10321
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13587

Query Match
Best Local Similarity 5.9%; Score 35.2; DB 4; Length 10321;
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 250 TGAGGAGGGATGTCCTATGTTAGAGATTCAGAGATAAGTTTCAGGAATGTAACTTATA 309
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7305 TGACGAGGAGGATGGGTGTGGAGACAGAGAGAGTTGGAGTGGAGAGAGAGGGA 7246
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GATTTTATACATACACAGAGAAATACGACTAGTGAGAAAGCTATTGCCATGTCCTCAAGCA 369
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7245 AAGGAGAGAGAGACAAAGAGATGTAGAGAGAGGAAAGGTTATGGGAGAGAGAGAC 7186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
370 AGAGATGATGAAGCCTTAATATGGAGCCAAA 401
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
7185 AGAGAGATGGGGTGGAAAGAGAGAGGGAGA 7154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(786431)
; OTHER INFORMATION: n = A,T,C or G
; US-09-751-389-3

Query Match
Best Local Similarity 5.9%; Score 35.2; DB 4; Length 786431;
Matches 94; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 142 GGATGACAAATTAGAAAGGAGGTTGAGTCAAGGAGTTTGAATGTTAGGGTGAAAGTTAC 201
Db 512627 GAAAGAGAAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512686

QY 202 TACTCACTCTGTAGGTTTAAAGGAAAGCTTTCAGTCCCAATGAGAGGGATG 261
Db 512687 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512746

QY 262 TGCCATGTTTAGAGATTTCAGAGATAAGTTTCAGGAATGTAACTATAGATTTTATACAT 321
Db 512747 GCGCAAGTTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512806

QY 322 ACACAGAGAAAT 333
Db 512807 ACAATGACAATT 512818

RESULT 15
US-09-949-016-16678
; Sequence 16678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16678
; LENGTH: 84525
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16678

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 84525;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 206 CAACTCTGTAGTTTAAAGGAAACGTTTGAGAACTTTCAGTCCCAATGAGGAGGGATGTGCC 265
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 CAGACCTGGAGTGTAAACGGTAATAATTCAAAGGTTATATTCGTTAAGGCCAAATGTGGC 554
QY 266 ATGTTTAGAGATTCAGAGATAAGTTTCAGGAAATGTAACCTTATAGATTTTATACATACAC 325
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555 ATTTTGAGGATTACTTTGTATTTTATGAGACTGTAGCTTGTGATTTAGGCTTTCCAA 614
QY 326 AGA 328
Db 615 AAA 617

Search completed: August 4, 2005, 14:40:33
Job time : 128.644 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 1070.28 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-1  
Perfect score: 600  
Sequence: 1 taggaataattagatcat.....ctggaatcattctaa 600

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues  
Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				Published Applications NA:*
1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:			
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:			
3:	/cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:			
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:			
5:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:			
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:			
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:			
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:			
9:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:			
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:			
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:			
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:			
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:			
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:			
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:			
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:			
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:			
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:			
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:			
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:			
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:			
22:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:			
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:			
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:			
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:			
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	600	9	US-09-899-276-1
2	600	100.0	11793	19	US-10-685-705-4
3	42.6	7.1	96597	17	US-10-085-117-112
4	42.2	7.0	408	21	US-10-275-333A-13
5	41.8	7.0	760	13	US-10-027-632-170737
6	41.8	7.0	760	17	US-10-027-632-170737
7	41.2	6.9	78268	13	US-10-087-192-742

8	39.8	6.6	59001	21	US-10-317-869A-4	Sequence 4, Appli
9	39.6	6.6	329	19	US-10-674-124A-22067	Sequence 22067, A
C 10	39.4	6.6	126974	19	US-10-304-019-12	Sequence 12, Appl
C 11	39.4	6.6	126974	24	US-11-013-608-12	Sequence 12, Appl
C 12	39.4	6.6	144035	13	US-10-087-192-322	Sequence 322, App
C 13	39.2	6.5	3918	13	US-10-027-632-263881	Sequence 263881,
C 14	39.2	6.5	3918	13	US-10-027-632-263881	Sequence 263881,
C 15	39.2	6.5	3918	17	US-10-027-632-263881	Sequence 263881,
C 16	39.2	6.5	3918	17	US-10-027-632-263881	Sequence 263881,
C 17	38.8	6.5	388	18	US-10-424-599-10647	Sequence 10647, A
C 18	38.8	6.5	1768	19	US-10-767-795-5448	Sequence 5448, Ap
C 19	38.4	6.4	8943	17	US-10-257-166-48	Sequence 48, Appl
20	38.2	6.4	52302	11	US-09-997-722-4	Sequence 4, Appli
21	37.6	6.3	520	14	US-10-123-155-144	Sequence 144, App
22	37.6	6.3	520	15	US-10-146-731-144	Sequence 144, App
23	37.6	6.3	520	15	US-10-140-472-144	Sequence 144, App
24	37.6	6.3	520	15	US-10-141-761-144	Sequence 144, App
25	37.6	6.3	520	16	US-10-142-885-144	Sequence 144, App
26	37.6	6.3	520	16	US-10-158-790-144	Sequence 144, App
27	37.6	6.3	520	17	US-10-137-871-144	Sequence 144, App
28	37.6	6.3	520	17	US-10-140-923-144	Sequence 144, App
29	37.6	6.3	520	17	US-10-141-756-144	Sequence 144, App
30	37.6	6.3	520	17	US-10-141-759-144	Sequence 144, App
31	37.6	6.3	520	17	US-10-140-805-144	Sequence 144, App
32	37.6	6.3	520	17	US-10-140-864-144	Sequence 144, App
33	37.6	6.3	520	18	US-10-142-426-144	Sequence 144, App
34	37.2	6.2	398287	19	US-10-741-601-5719	Sequence 5719, Ap
35	37.2	6.2	398287	21	US-10-741-600-17839	Sequence 17839, A
36	37	6.2	2545	17	US-10-115-831-155	Sequence 155, App
37	37	6.2	7261	18	US-10-221-714A-391	Sequence 391, App
38	37	6.2	401616	13	US-10-087-192-1630	Sequence 1630, Ap
C 39	36.8	6.1	638	13	US-10-027-632-190233	Sequence 190233,
C 40	36.8	6.1	638	17	US-10-027-632-190233	Sequence 190233,
C 41	36.8	6.1	1894	20	US-10-425-115-171730	Sequence 171730,
C 42	36.6	6.1	201143	18	US-10-240-425-1099	Sequence 1099, Ap
C 43	36.6	6.1	3673778	16	US-10-312-841-2	Sequence 2, Appli
C 44	36.4	6.1	901	13	US-10-027-632-161407	Sequence 161407,
45	36.4	6.1	901	17	US-10-027-632-161407	Sequence 161407,

ALIGNMENTS

RESULT 1  
US-09-899-276-1  
; Sequence 1, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Deltus, HaJo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899,276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-1

Query Match 100.0%; Score 600; DB 9; Length 600;  
Best Local Similarity 100.0%; Pred. No. 9.6e-167;  
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TAGGAAAATTATAGGATCATTAAAGAAAGGAAAGAAAGTGGAGCAAAATACCTGGAGG	60
Db	1	TAGGAAAATTATAGGATCATTAAAGAAAGGAAAGAAAGTGGAGCAAAATACCTGGAGG	60
QY	61	TAGAAATGGTGATGTGTATCATCAAGCAGGGAGAAACCAANTGAACGATGCGAATT	120
Db	61	TAGAAATGGTGATGTGTATCATCAAGCAGGGAGAAACCAANTGAACGATGCGAATT	120
QY	121	CGGGCCCAACCAATGTCAAGGGATGACAAATTAGAAAGAAAGAGTTGAGTCAAGGATTTG	180
Db	121	CGGGCCCAACCAATGTCAAGGGATGACAAATTAGAAAGAAAGAGTTGAGTCAAGGATTTG	180
QY	181	AATGTTAGGGTGAAAAAGTTACTATCAACTCTGTAGGTTAAAGGAAACGTTGAGAATCT	240
Db	181	AATGTTAGGGTGAAAAAGTTACTATCAACTCTGTAGGTTAAAGGAAACGTTGAGAATCT	240
QY	241	TCAGTCCAAATGAGGAGGATGTCCTCATGTTTATAGATTCAGAGATAAGTTTCAGAAATG	300
Db	241	TCAGTCCAAATGAGGAGGATGTCCTCATGTTTATAGATTCAGAGATAAGTTTCAGAAATG	300
QY	301	TAACTTATAGATTTTATACATACACAGAGAAATACGACTAGTGAGAACCTATGCCCATG	360
Db	301	TAACTTATAGATTTTATACATACACAGAGAAATACGACTAGTGAGAACCTATGCCCATG	360
QY	361	GTCCAAGCAGAGATGATGAAGGCGCTAAATATGGAGCCAAAGAGGAGCAATGAAGAATG	420
Db	361	GTCCAAGCAGAGATGATGAAGGCGCTAAATATGGAGCCAAAGAGGAGCAATGAAGAATG	420
QY	421	AGCCATGACGGTGAAATGCTGCAATGTTGTAATGGAGGAGAAAGACCTGTGACTTTCAGA	480
Db	421	AGCCATGACGGTGAAATGCTGCAATGTTGTAATGGAGGAGAAAGACCTGTGACTTTCAGA	480
QY	481	TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGAGCTTCCTGAAACCAAGATGT	540
Db	481	TATGAAAACCTCATCTTCAACCCACATTTTAAAGGGGAGCTTCCTGAAACCAAGATGT	540
QY	541	GTTTCCTCCATTACTATACCCCATCCCAATCTCAGGCACCTCGAATCATCCTTTAAA	600
Db	541	GTTTCCTCCATTACTATACCCCATCCCAATCTCAGGCACCTCGAATCATCCTTTAAA	600

RESULT 2  
 US-10-685-705-4  
 ; Sequence 4, Application US/10685705  
 ; Publication No. US2004017387A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Kentucky Research Foundation  
 ; APPLICANT: JAVAKRISHNA, Ambati  
 ; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
 ; TITLE OF INVENTION: Degeneration  
 ; FILE REFERENCE: 050229-0415  
 ; CURRENT APPLICATION NUMBER: US/10/685,705  
 ; CURRENT FILING DATE: 2003-10-16  
 ; PRIOR APPLICATION NUMBER: 60/422,096  
 ; PRIOR FILING DATE: 2002-10-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 4  
 ; LENGTH: 11793  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-685-705-4

Query Match	100.0%;	Score	600;	DB	19;	Length	11793;	
Best Local Similarity	100.0%;	Pred. No.	3.7e-166;					
Matches	600;	Conservative	0;	Mismatches	0;	Indels	0; Gaps	0;

QY 1 TAGGAAAATTATAGGATCATTAAAGAAAGGAAAGAGTGGAGCAAAATACCTGGAGG 60  
 Db 9979 TAGGAAAATTATAGGATCATTAAAGAAAGGAAAGAGTGGAGCAAAATACCTGGAGG 10038  
 QY 61 TAGAAATGGTGATGTGTATCATCAAGCAGGGAGAAACCAANTGAACGATGCGAATT 120



σδφππ: T-39/7-668-60-8η

```

? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: fastseq for Windows Version 4.0
? SEQ ID NO 170737
? LENGTH: 760
? TYPE: DNA
? ORGANISM: Human
? RS-10-027-632-170737

```

	Query Match	7.08;	Score 41.8;	DB 17;	Length 760;
	Best Local Similarity	50.2%;	Prod. NO. 0.13;		
	Matches 103;	Conservative	0;	Mismatches 102;	Indels 0;
				Gaps	0;
QY	32	RAGGAAGAGTGGGAGCAAAATACCTGAGAGGTAGAAAATGGTGTGATGTGTACTCATCAGCAG	91		
Db	556	RAGAAGAAGAGCGGAGCGGAGGAGGAAGGAAGAAGGAAGGAAGGAAGGAAGGAAGGAAG	615		
QY	92	GGAGAAAAACCAATGAACCGAGATCGGAATTCGGGCCCCACACCAATGTCAAGGGATGACCAAT	151		
Db	616	CGAGAGAAGGAAGGAAGGAAGGAGGAGGGAGGGGAGGGAATATAGNAGGAGGCGCAGAAAGA	675		
QY	152	TAGAAGGAAGGTTGAGTCAAGGGAATTTGAATGTTAGGGGTGAAGATTACTACTCAACTC	211		
Db	676	GAGAGAGGAGGAGGAGGAAGCAATTGAAATTTTGGACTACAGAGAAGAAAATAACAG	735		
QY	212	TGTAGGTTAAAAAGGAAACGTTTCAG	236		
Db	736	GAATGAGTTTAAGTGAGGAGGTGGGA	760		

```

RESULT 7
; US-10-087-192-742
; Sequence 742, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 742
; LENGTH: 78268
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-087-192-742

```

	Query Match	Best Local Similarity	6.9%;	Score 41.2;	DB 13;	Length 78268;
	Matches	Conservative	62.7%;	Prod No. 1.6;	Mismatches 38;	Indels 0; Gaps 0
QY	26	AAGGAGAAAGGAGTAGTGGAGCAATACCTCGGAGGTAGAAATCGTGATGATGTGTACATC	85			
DB	9409	AGGCACAAGAAACAGCAGGTACAAAGCTCGGAGTTCGGAACCTTAATGTGTTCACA	94			
QY	86	ARCGAGGGAGAAACCAATGAACAGATCGGAATTCGGGCC	127			
DB	9469	GAACGAAGAAGAACCAATGTACCTGGAGAGCAGCCAGGCC	9510			

```

: RESULT 5
: US-10-027-632-170737
: Sequence 170737, Application US/10027632
: Publication NO. US20020198371A1
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: POLYMORPHISMS IN THE HUMAN GENOME
: TITLE OF INVENTION: Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIORITY APPLICATION NUMBER: US 60/218,006

```

```

, PRIOR FILING DATE: 2000-07-12
, PRIOR APPLICATION NUMBER: US 60/198,676
, PRIOR FILING DATE: 2000-04-20
, PRIOR APPLICATION NUMBER: US 60/193,483
, PRIOR FILING DATE: 2000-03-29
, PRIOR APPLICATION NUMBER: US 60/195,218
, PRIOR FILING DATE: 2000-02-24
, PRIOR APPLICATION NUMBER: US 60/167,363
, PRIOR FILING DATE: 1999-11-23
, PRIOR APPLICATION NUMBER: US 60/156,358
, PRIOR FILING DATE: 1999-09-28
, PRIOR APPLICATION NUMBER: US 60/146,002
, PRIOR FILING DATE: 1999-08-09
, NUMBER OF SEQ ID NOS: 325720
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 170737
, LENGTH: 760
, TYPE: DNA
, ORGANISM: Human
, NC_10-027,632-170737

```

[illegible]

```

; Publication No. US20030204075A9
;
; GENERAL INFORMATION:
;
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
;
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
;
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
;
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
;

```

## RESULT 8

US-10-317-869A-4  
; Sequence 4, Application US/10317869A  
; Publication No. US20050101000A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Kenneth W. Doble  
; TITLE OF INVENTION: MODULATION OF PHOSPHODIESTERASE 4B EXPRESSION  
; FILE REFERENCE: RTS-0429  
; CURRENT APPLICATION NUMBER: US/10/317,869A  
; CURRENT FILING DATE: 2002-12-11  
; NUMBER OF SEQ ID NOS: 113  
; SEQ ID NO 4  
; LENGTH: 599001  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 55607-55706, 360227-360326, 363103-363202, 363428, 366093-366192,  
; LOCATION: 368528, 447532, 452653-452752, 457192, 457207-457306, 461808-461907,  
; LOCATION: 461921, 461928, 461942  
; OTHER INFORMATION: n = A, T, C or G  
US-10-317-869A-4

Query Match 6.6%; Score 39.8; DB 21; Length 599001;

Best Local Similarity 48.5%; Pred. No. 10;

Matches 144; Conservative 0; Mismatches 147; Indels 6; Gaps 1;

QY	196	AGTTACTCACTCTGTAGGTTAAAGGAAACGTTTGAGAACTTCAGTCCTCCATGAGGA	255
DB	512227	AGGTGCTGTAAGCTATTGGATTATATTAAAGCAAAATGGGCAGCCACTGAAAGAAATG	512286
QY	256	GGGATGTCGCATGTTTAGAGATTCAGAGATAAGTTTCAGAAATGTAAC-----TTATA	309
DB	512287	GAGAAGAGAAATGTTAT	512346
QY	310	GATTTTATACATACAGAGAAATACGGACTAGTGAGAGCTATTGCCATGTGTCACAGCA	369
DB	512347	TGTGTAATAAATGGGATTGGAAGAGAGGATATGAGGGGAAATGTTCAAGAACTCTGAGCA	512406
QY	370	AGAGATGATGAAGCCCTTAATATATGGAGCCCAAGAGCAGCAATGAAGATGAGCCATGCA	429
DB	512407	AGAGGTGATGCTTGGCTTAATCAAGATGTTAGCAGGAGGAAGTGGAGGAACTGGACCGAT	512466
QY	430	GGGTGAAATGCTGCATGTTGTTAAATGGAGAGAAACCTGTGACTTCAGATATGAA	486
DB	512467	GTGTAATATATCTGTAGGTGAATAGATAGATGATGCTGATAAATTAGTGGGA	512523

## RESULT 9

US-10-674-124A-22067  
; Sequence 22067, Application US/10674124A  
; Publication No. US20040197797A1  
; GENERAL INFORMATION:  
; APPLICANT: INOKO, Hidetoshi  
; APPLICANT: TAMURA, Gen  
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
; FILE REFERENCE: ORIN-003CIP  
; CURRENT APPLICATION NUMBER: US/10/674,124A  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: 10/257,511  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: PCT/JP00/07621  
; PRIOR FILING DATE: 2000-10-30  
; PRIOR APPLICATION NUMBER: JP2000-112699  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: JP2002-327516  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: JP2002-383869  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 27110  
; SEQ ID NO 22067

; LENGTH: 329

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: HUMUT8173B

; FEATURE:

; OTHER INFORMATION: Located on chromosome 16

; FEATURE:

; OTHER INFORMATION: Distance between a terminus base of telomere on

; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base

; OTHER INFORMATION: sequence : 28899985

; FEATURE:

; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of

; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and

; OTHER INFORMATION: 5'-terminus of this base sequence : 216681

US-10-674-124A-22067

Query Match 6.6%; Score 39.6; DB 19; Length 329;

Best Local Similarity 51.7%; Pred. No. 0.4;

Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY	22	AAGAAAGGAGAGAGAGTGGGAGCAATACCTGGAGGTAGAAATGGTGATGATGTGTA	81
DB	29	AAGAAAGGAGAGAGAGGAGGAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	88
QY	82	CATCAAGCAGGAGGAGAAACCAATGAACCAAGATCGCAATTCGGGCCCCACACCAATGTCAAG	141
DB	89	AAGAAAGGAG	148
QY	142	GGATGACAAATTAGAAAGGAGTTGAGTCAAGGGATTTGAAATGTTAGGGTGAAA	195
DB	149	CGAGGAG	202

## RESULT 10

US-10-304-019-12/c  
; Sequence 12, Application US/10304019  
; Publication No. US20040102622A1  
; GENERAL INFORMATION:  
; APPLICANT: Nicholas M. Dean  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Kenneth W. Doble  
; TITLE OF INVENTION: MODULATION OF HEPATOCYTE GROWTH FACTOR RECEPTOR EXPRESSION  
; FILE REFERENCE: PTS-0043  
; CURRENT APPLICATION NUMBER: US/10/304,019  
; CURRENT FILING DATE: 2002-11-23  
; NUMBER OF SEQ ID NOS: 147  
; SEQ ID NO 12  
; LENGTH: 126974  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
US-10-304-019-12

Query Match 6.6%; Score 39.4; DB 19; Length 126974;

Best Local Similarity 50.8%; Pred. No. 6.8;

Matches 94; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY	287	AGTTTCAGAAATGTAACTTATAGATTTTATACATACAGAGAAATACGGAATAGTAG	346
DB	88257	AATTACACAAAATAAATATTTTCCCTCTGAAAAAATAACAAAGAAACCCGAATTAAC	88198
QY	347	AGCTATTGCCATGTCCTCAAGCAAGAGATGTAAGGGCTAATATGGAGCCAAAGAGGC	406
DB	88197	TTGTTGCTAGACTGATAGATTTAAACGGAACAGGAGATTTGATGCTGGTCCATAGAT	88138
QY	407	AGCAATGAAGATGAGCCATGCGAGGTGAAATGCTGCATGTTGTAATGGAGGAGAGA	466
DB	88137	ATCAATTCAAAGAGAGTCTGGCATAAAGAAAGAGCTGTGTGTAAGCAGCAGCAGCA	88078
QY	467	CCTGT 471	
DB	88077	CATAT 88073	

Tue Aug 9 17:01:24 2005

```

; Sequence 322, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 144035
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-087-192-322

Query Match      6.6%; Score 39.4; DB 13; Length 144035;
Best Local Similarity 50.8%; Pred. No. 7.2; Indels 0; Gaps 0;
Matches 94; Conservative

QY 287 AGTTTCAGGAATGTAACCTTATAGATTTTATACATACACAGAGAAATACGGACTAGTGAG 346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96760 AATTACACAAATTAATATTTCCCTCTGAAAAAATACAAAGAAAACCGAATTAACCTA 96701

QY 347 AAGCTATTGCCATGCTCCAGCAGAGATGATGAGGCCTAAATATGAGCCAAAGAGGC 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96700 TTGTTGCTAGACTGATAGATTAACGGAAAACAAAGCAGATGAACTGGTCCATGAGAT 96641

QY 407 AGCAATGAAGATGAGCAGCATGCGGTGAAATGCTGCAATGTTGTAATGAGGAGAAAGA 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96640 ATCAATTCAAAGAGAGCTGCGCATTAAGAAAGCTGTGTGTTAAAGCAGGAGGACAAAA 96581

QY 467 CCTGT 471
    |||||
Db 96580 CATAT 96576
    |||||

```

```

RESULT 13
US-10-027-632-263881/c
; Sequence 263881, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263881
; LENGTH: 3918
; TYPE: DNA

```

```

RESULT 11
US-11-013-608-12/c
; Sequence 12, Application US/11013608
; Publication No. US20050153925A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Brett P. Monia
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan M. Freier
; APPLICANT: Brenda F. Baker
; APPLICANT: Donna T. Ward
; APPLICANT: William A. Gaarde
; APPLICANT: Lex M. Cowsett
; APPLICANT: Andrew T. Watt
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO TRANSFERASES
; FILE REFERENCE: BNDL-0009US:P1
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/11/013,608
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 10/303,329
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 10/298,994
; PRIOR FILING DATE: 2002-11-16
; PRIOR APPLICATION NUMBER: 10/300,642
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 10/298,123
; PRIOR FILING DATE: 2002-11-16
; PRIOR APPLICATION NUMBER: 10/303,541
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 10/303,588
; PRIOR FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: 10/293,863
; PRIOR FILING DATE: 2002-11-11
; PRIOR APPLICATION NUMBER: 10/298,953
; PRIOR FILING DATE: 2002-11-16
; PRIOR APPLICATION NUMBER: 10/167,034
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 10/175,627
; PRIOR FILING DATE: 2002-06-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 147
; SEQ ID NO 12
; LENGTH: 126974
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
US-11-013-608-12

Query Match      6.6%; Score 39.4; DB 24; Length 126974;
Best Local Similarity 50.8%; Pred. No. 6.8; Indels 0; Gaps 0;
Matches 94; Conservative

QY 287 AGTTTCAGGAATGTAACCTTATAGATTTTATACATACACAGAGAAATACGGACTAGTGAG 346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88257 AATTACACAAATAAATATTTCCCTCTGAAAAAATACAAAGAAAACCGAATTAACCTA 88198

QY 347 AAGCTATTGCCATGCTCCAGCAGAGATGATGAGGCCTAAATATGAGCCAAAGAGGC 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88197 TTGTTGCTAGACTGATAGATTAACGGAAAACAAAGCAGATGAACTGGGTCCAATGAGAT 88138

QY 407 AGCAATGAAGATGAGCAGCATGCGGTGAAATCCTCATGTTGTAATGAGGAGAAAGA 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88137 ATCAATTCAAAGAGAGCTGCGCATTAAGAAAAGCTGTGTGTAAGCAGCAGGACAAAAA 88078

QY 467 CCTGT 471
    |||||
Db 88077 CATAT 88073
    |||||

```

RESULT 12  
US-10-087-192-322/c

```
; ORGANISM: Human
US-10-027-632-263881

Query Match
Best Local Similarity 6.5%; Score 39.2; DB 13; Length 3918;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 324 ACAGAGAAATACGACTAGTGAAGCTATTGGCATGGTCCAAAGCAAGAGATGATGAAGG 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 ATAGAGCGCAGGAGACAAATTAGAGCCATTGCAATCATCCAAAGTGAGAGGTGCTGAATG 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 384 CCTAAATATGAGCCCAAGAGCGAGCAATGAA 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CCTAAATTAGGAAAGCAGTAGCAATATAGAA 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 14
US-10-027-632-263882/c
; Sequence 263882, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263882
; LENGTH: 3918
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-263882
```

```
Query Match
Best Local Similarity 6.5%; Score 39.2; DB 13; Length 3918;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 324 ACAGAGAAATACGACTAGTGAAGCTATTGGCATGGTCCAAAGCAAGAGATGATGAAGG 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 ATAGAGCGCAGGAGACAAATTAGAGCCATTGCAATCATCCAAAGTGAGAGGTGCTGAATG 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 384 CCTAAATATGAGCCCAAGAGCGAGCAATGAA 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CCTAAATTAGGAAAGCAGTAGCAATATAGAA 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 15
US-10-027-632-263881/c
; Sequence 263881, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263881
; LENGTH: 3918
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-263881
```

```
Query Match
Best Local Similarity 6.5%; Score 39.2; DB 17; Length 3918;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 324 ACAGAGAAATACGACTAGTGAAGCTATTGGCATGGTCCAAAGCAAGAGATGATGAAGG 383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 ATAGAGCGCAGGAGACAAATTAGAGCCATTGCAATCATCCAAAGTGAGAGGTGCTGAATG 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 384 CCTAAATATGAGCCCAAGAGCGAGCAATGAA 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 CCTAAATTAGGAAAGCAGTAGCAATATAGAA 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: August 4, 2005, 15:54:10  
Job time : 1076.28 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 803.887 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-2

Perfect score: 200

Sequence: 1 tgcagtaactattttcc.....aattgtttctctgtac 200

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.ste.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	200	100.0	200	6	AX343329
2	198.4	99.2	11793	9	HSV18933
3	196.8	98.4	360	6	CQ069701 Sequence
4	196.8	98.4	360	6	CQ096880 Sequence
5	196.8	98.4	360	6	CQ135681 Sequence
6	196.8	98.4	360	6	CQ174058 Sequence
7	196.8	98.4	360	6	CQ218977 Sequence
8	196.8	98.4	360	6	CQ257544 Sequence
9	196.8	98.4	360	6	CQ294784 Sequence
10	196.8	98.4	360	6	CQ331440 Sequence
11	196.8	98.4	508	6	CQ861661 Sequence
12	196.8	98.4	647	6	AR567996 Sequence
13	196.8	98.4	647	6	AX698741 Sequence
14	196.8	98.4	725	6	AR337874 Sequence
15	196.8	98.4	725	6	AR380799 Sequence
16	196.8	98.4	725	6	AX774742 Sequence
17	196.8	98.4	725	9	HSWCPI
18	196.8	98.4	729	6	CQ724136 Sequence
19	196.8	98.4	739	6	AX577974 Sequence

20	196.8	98.4	739	9	S71513	S71513 monocyte ch
21	196.8	98.4	741	6	A17786	A17786 MCP-1 mRNA
22	196.8	98.4	741	6	AR094465	AR094465 Sequence
23	196.8	98.4	741	6	E05611	E05611 CDNA encodi
24	196.8	98.4	741	6	AR380620	AR380620 Sequence
25	196.8	98.4	741	9	HUMMCAF	M24545 Human monocy
26	196.8	98.4	743	9	BC009716	BC009716 Homo sapi
27	196.8	98.4	757	6	CQ831706	CQ831706 Sequence
28	196.8	98.4	1712	6	BD195644	BD195644 70 human
29	196.8	98.4	1712	6	CQ775689	CQ775689 Sequence
30	196.8	98.4	1712	6	AR352699	AR352699 Sequence
31	196.8	98.4	1822	6	BD195643	BD195643 70 human
32	196.8	98.4	1822	6	CQ775688	CQ775688 Sequence
33	196.8	98.4	1822	6	AR352698	AR352698 Sequence
34	196.8	98.4	2776	9	HUMMCHEMP	M37719 Human monocy
35	196.8	98.4	9174	9	AF519531	AF519531 Homo sapi
36	196.8	98.4	147416	9	AC005549	AC005549 Homo sapi
37	195.2	97.6	661	6	AR559337	AR559337 Sequence
38	195.2	97.6	661	6	BD080551	BD080551 Chemokine
39	195.2	97.6	661	9	S69738	S69738 MCP-1 monocy
40	195.2	97.6	2243	9	HSJEPR	X60001 H.sapiens g
41	184.8	92.4	880	11	G06615	G06615 human STS W
42	184.8	92.4	1011	6	A37281	A37281 Sequence 21
43	184.8	92.4	1011	6	CQ861415	CQ861415 Sequence
44	184.8	92.4	1011	6	AX331515	AX331515 Sequence
45	184.8	92.4	1011	6	AX663653	AX663653 Sequence

#### ALIGNMENTS

RESULT 1	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
LOCUS	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
DEFINITION	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
ACCESSION	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
VERSION	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
KEYWORDS	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
SOURCE	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
ORGANISM	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
REFERENCE	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
AUTHORS	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
TITLE	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
JOURNAL	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002

FEATURES	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
source	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
Location/Qualifiers	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
1..200	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
/organism="Homo sapiens"	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
/mol_type="unassigned DNA"	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002
/db_xref="taxon:9606"	AX343329.1	GI:18491679	200 bp	DNA	linear	PAT 01-FEB-2002

ORIGIN	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
Query Match	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
Best Local Similarity	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	AX343329	Sequence 2 from Patent EP1170372.	200 bp	DNA	linear	PAT 01-FEB-2002
Qy	1	TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACTTTGTTTATTTATTTATTAATG 60				
Db	1	TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACTTTGTTTATTTATTTATTAATG 60				
Qy	61	AATTTTGTGTTGATGTGAACATTATGCTTAACTAATGTTATTTCTTTATTTAAGTTA 120				
Db	61	AATTTTGTGTTGATGTGAACATTATGCTTAACTAATGTTATTTCTTTATTTAAGTTA 120				
Qy	121	TGATGTTTTTAAAGTTTATCTTTTCATGGTACTAGTGTGTTTTTTAGATACAGACTTTGGGGA 180				
Db	121	TGATGTTTTTAAAGTTTATCTTTTCATGGTACTAGTGTGTTTTTTAGATACAGACTTTGGGGA 180				
Qy	181	AATTCCTTTTCTCTCTGTAC 200				

Tue Aug 9 17:01:25 2005

```

Db      181 AATTGCTTTTCCTCTGTAC 200
|||||
HSY18933      11793 bp      DNA      linear      PRI 18-OCT-2000
LOCUS      Homo sapiens MCP-1 gene and enhancer region.
DEFINITION      Y18933
ACCESSION      Y18933.1 GI:10933860
VERSION      MCP-1 gene; monocyte chemoattractant protein-1.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
REFERENCE      1
AUTHORS      Finzer, P., Soto, U., Delius, H., Patzelt, A., Coy, J.F., Pousetka, A.,
zur Hausen, H. and Roel, F.
TITLE      Differential transcriptional regulation of the
monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and
non-tumorigenic HPV 18 positive cells: the role of the chromatin
structure and AP-1 composition
Oncogene 19 (29), 3235-3244 (2000)
JOURNAL      Oncogene 19 (29), 3235-3244 (2000)
MEDLINE      20374005
PubMed      10918580
REFERENCE      2 (bases 1 to 11793)
AUTHORS      Roel, F.
TITLE      Direct Submission
Submitted (15-APR-1999) F. Roel, Applied Tumor Virology, DKPZ
(German Cancer Research Institute), Im Neuenheimer Feld 506, 69120
Heidelberg, FRG
FEATURES      source
              1. 11793
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="17"
              4906..9073
              /gene="MCP-1"
              4906..4915
              /gene="MCP-1"
              4935..4944
              /gene="MCP-1"
              7513..7517
              /gene="MCP-1"
              join(<7550..7685,8468..8585,8968..>9073)
              /gene="MCP-1"
              7550..7685
              /gene="MCP-1"
              /number=1
              join(7610..7685,8468..8585,8968..9073)
              /gene="MCP-1"
              /codon_start=1
              /product="monocyte chemoattractant protein-1"
              /protein_id="CAC14049.1"
              /db_xref="GI:10933861"
              /db_xref="GGA:PI3500"
              /db_xref="UniProt/Swiss-Prot:P13500"
              /translation="MKVSAALLCLLLIATFI PQGLAQPDALNAPVTCYNTNRKIS
VQLASVRRITSSKCPKEAVIFKTIIVAKEICADPKQKWQSDMDHLDKQTQPKT"
              7686..8467
              /gene="MCP-1"
              /number=1
              8468..8585
              /gene="MCP-1"
              /number=2
              8586..8967
              /gene="MCP-1"
              /number=2
              8968..>9073
              /gene="MCP-1"
              /number=3
              intron
              exon
              intron
              exon
              ORIGIN

```

```

Query Match      99.2%; Score 198.4; DB 9; Length 11793;
Best Local Similarity 99.5%; Pred. No. 1e-26; Indels 0; Gaps 0;
Matches 199; Conservative
QY      1 TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACCTTGTATTTATTTATTAATG 60
Db      9099 TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACCTTGTATTTATTTATTAATG 9158
QY      61 AATTTTGGTTTGTGATGGAACATATGCCCTTAAGTAATGTTAATCTTTATTTAAGTTA 120
Db      9159 AATTTTGGTTTGTGATGGAACATATGCCCTTAAGTAATGTTAATCTTTATTTAAGTTA 9218
QY      121 TTGATGTTTAAAGTTATATCTTTTCATGGTACTAGTCTTTTATAGATACAGAGACTTTGGGGA 180
Db      9219 TTGATGTTTAAAGTTATATCTTTTCATGGTACTAGTCTTTTATAGATACAGAGACTTTGGGGA 9278
QY      181 AATTGCTTTTCCTCTCTGTAC 200
Db      9279 AATTGCTTTTCCTCTCTGTAC 9298
RESULT 3
CQ069701/c      CQ069701      360 bp      DNA      linear      PAT 20-JAN-2004
LOCUS      Sequence 5501 from Patent WO0157278.
DEFINITION      CQ069701
ACCESSION      CQ069701
VERSION      CQ069701.1 GI:41039570
KEYWORDS      Homo sapiens (human)
ORGANISM      Homo sapiens
SOURCE      Homo sapiens
REFERENCE      1
AUTHORS      Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE      Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human hela cells or other human
cervical epithelial cells
Patent: WO 0157278-A 5501 09-AUG-2001;
JOURNAL      Aecomica, Inc. (US)
FEATURES      Location/Qualifiers
              1..360
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              /note="MAP TO AC005549.1-EXPRESSED IN HELA, SIGNAL = 2.1"
ORIGIN
Query Match      98.4%; Score 196.8; DB 6; Length 360;
Best Local Similarity 99.0%; Pred. No. 4.3e-26; Indels 0; Gaps 0;
Matches 198; Conservative
QY      1 TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACCTTGTATTTATTTATTAATG 60
Db      227 TGCAGCTAACTATTATTTCCCTAGCTTTCCCGACACACCTTGTATTTATTTATTAATG 168
QY      61 AATTTTGGTTTGTGATGGAACATATGCCCTTAAGTAATGTTAATCTTTATTTAAGTTA 120
Db      167 AATTTTGGTTTGTGATGGAACATATGCCCTTAAGTAATGTTAATCTTTATTTAAGTTA 108
QY      121 TTGATGTTTAAAGTTATATCTTTTCATGGTACTAGTCTTTTATAGATACAGAGACTTTGGGGA 180
Db      107 TTGATGTTTAAAGTTATATCTTTTCATGGTACTAGTCTTTTATAGATACAGAGACTTTGGGGA 48
QY      181 AATTGCTTTTCCTCTCTGTAC 200
Db      47 AATTGCTTTTCCTCTCTGTAC 28
RESULT 4
CQ096880/c      CQ096880      360 bp      DNA      linear      PAT 21-JAN-2004
LOCUS      Sequence 5739 from Patent WO0157272.
DEFINITION

```

ACCESSION CQ096880  
VERSION CQ096880.1 GI:41065906  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human placenta  
JOURNAL Patent: WO 0157272-A 5739 09-AUG-2001;  
Aeomica, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN PLACENTA, SIGNAL = 2.9"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 360;  
Best Local Similarity 99.0%; Pred. No. 4.3e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 60  
DB 227 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 168

QY 61 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 120  
DB 167 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 108

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 180  
DB 107 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 48

QY 181 AATTGCTTTTCCCTCTGTAC 200  
DB 47 AATTGCTTTTCCCTCTGTAC 28

RESULT 5  
LOCUS CQ135681 360 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 5703 from Patent WO0157276.  
ACCESSION CQ135681  
VERSION CQ135681.1 GI:41093047  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human bone marrow  
JOURNAL Patent: WO 0157276-A 5703 09-AUG-2001;  
Aeomica, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN BONE MARROW, SIGNAL = 2.4"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 360;  
Best Local Similarity 99.0%; Pred. No. 4.3e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 60  
DB 227 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 168

QY 61 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 120  
DB 167 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 108

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 180  
DB 107 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 48

QY 181 AATTGCTTTTCCCTCTGTAC 200  
DB 47 AATTGCTTTTCCCTCTGTAC 28

RESULT 6  
LOCUS CQ174058 360 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 5454 from Patent WO0157274.  
ACCESSION CQ174058  
VERSION CQ174058.1 GI:41168794  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human heart  
JOURNAL Patent: WO 0157274-A 5454 09-AUG-2001;  
Aeomica, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN HEART, SIGNAL = 2.6"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 360;  
Best Local Similarity 99.0%; Pred. No. 4.3e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 60  
DB 227 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 168

QY 61 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 120  
DB 167 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 108

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 180  
DB 107 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 48

QY 181 AATTGCTTTTCCCTCTGTAC 200  
DB 47 AATTGCTTTTCCCTCTGTAC 28

RESULT 7  
LOCUS CQ218977 360 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 5816 from Patent WO0157273.  
ACCESSION CQ218977  
VERSION CQ218977.1 GI:41200544  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human bone marrow  
JOURNAL Patent: WO 0157273-A 5816 09-AUG-2001;  
Aeomica, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN BONE MARROW, SIGNAL = 2.4"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 360;  
Best Local Similarity 99.0%; Pred. No. 4.3e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 60  
DB 227 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTATAATG 168

QY 61 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 120  
DB 167 AATTTTGTGTGATGTAACATTAATGCTTAAGTAATGTAATCTTATTAAAGTTA 108

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 180  
DB 107 TTGATGTTTAAAGTTTATCTTTTCATGCTAGTACTGTTTATTTAGATACAGAGACTTGGGGA 48

QY 181 AATTGCTTTTCCCTCTGTAC 200  
DB 47 AATTGCTTTTCCCTCTGTAC 28

Tue Aug 9 17:01:25 2005

```

REFERENCE
AUTHORS
TITLE
1
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 5816 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .360
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC005549.1-EXPRESSED IN ADULT LIVER, SIGNAL
= 2.5"
ORIGIN
Query Match 98.4%; Score 196.8; DB 6; Length 360;
Best Local Similarity 99.0%; Pred. No. 4.3e-26;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATTATAATG 60
Db 227 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATTATAATG 168
QY 61 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 120
Db 167 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 108
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
Db 107 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 48
QY 181 AATTGCTTTTCCTCTTGAC 200
Db 47 AATTGCTTTTCCTCTTGAC 28
RESULT 9
CO257544/c
LOCUS
DEFINITION
Sequence 5805 from Patent WO0157277.
ACCESSION
CO257544
VERSION
CO257544.1 GI:41230024
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human fetal liver
Patent: WO 0186003-A 5889 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .360
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC005549.1-EXPRESSED IN LUNG, SIGNAL = 2.5"
ORIGIN
Query Match 98.4%; Score 196.8; DB 6; Length 360;
Best Local Similarity 99.0%; Pred. No. 4.3e-26;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATAATG 60
Db 227 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATAATG 168
QY 61 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 120
Db 167 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 108
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
Db 107 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 48
QY 181 AATTGCTTTTCCTCTTGAC 200
Db 47 AATTGCTTTTCCTCTTGAC 28
RESULT 10
CO257544/c
LOCUS
DEFINITION
Sequence 5534 from Patent WO0157275.
ACCESSION
CO257544
VERSION
CO257544.1 GI:41280007
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human fetal liver
Patent: WO 0157277-A 5805 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .360
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC005549.1-EXPRESSED IN FETAL LIVER, SIGNAL
= 3.7"
ORIGIN
Query Match 98.4%; Score 196.8; DB 6; Length 360;
Best Local Similarity 99.0%; Pred. No. 4.3e-26;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATAATG 60
Db 227 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTTCCCGACACCTTTTATTATAATG 168
QY 61 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 120
Db 167 AATTTTGTGTTGATGTGAACATTTATGCCCTTAAGTAAATCTTATTAAAGTTA 108
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
Db 107 TTGATGTTTAAAGTTATCTTTTCATGCTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 48
QY 181 AATTGCTTTTCCTCTTGAC 200
Db 47 AATTGCTTTTCCTCTTGAC 28

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human brain  
JOURNAL Patent: WO 0157275-A 5534 09-AUG-2001;  
Neomica, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN BRAIN, SIGNAL = 2.5"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 360;  
Best Local Similarity 99.0%; Pred. No. 4.3e-26;  
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 60  
|||||  
Db 227 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 168  
|||||

QY 61 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 120  
|||||  
Db 167 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 108  
|||||

QY 121 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 180  
|||||  
Db 107 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 48  
|||||

QY 181 AATTGCTTTTCCCTCTGTAC 200  
|||||  
Db 47 AATTGCTTTTCCCTCTGTGAAC 28  
|||||

RESULT 11  
LOCUS Q0861661 508 bp DNA linear PAT 10-SEP-2004  
DEFINITION Sequence 294 from Patent WO2004072265.  
ACCESSION Q0861661  
VERSION Q0861661.1 GI:51982650  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Burczynski, M., Twine, N., Dörner, A.J. and Trepicchio, W.L.  
TITLE METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO /i  
JOURNAL Patent: WO 2004072265-A 294 26-AUG-2004;  
Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US); Dörner, Andrew J. (US); Trepicchio, William L. (US)  
FEATURES Location/Qualifiers  
source 1..508  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
misc\_feature 33..199  
/notes="Each #n# represents a nucleotide selected from a, t, g or c, or contains no nucleotide."

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 508;  
Best Local Similarity 99.0%; Pred. No. 4e-26;  
Matches 199; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 60  
|||||  
Db 256 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 315  
|||||

QY 61 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 120  
|||||

Db 316 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 375  
|||||

QY 121 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 180  
|||||

Db 376 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 435  
|||||

QY 181 AATTGCTTTTCCCTCTGTAC 200  
|||||

Db 436 AATTGCTTTTCCCTCTGTGAAC 455  
|||||

RESULT 12  
LOCUS AR567996 647 bp RNA linear PAT 08-OCT-2004  
DEFINITION Sequence 72 from patent US 6780982.  
ACCESSION AR567996  
VERSION AR567996.1 GI:53986227  
KEYWORDS SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 647)  
AUTHORS Lyanichev, V., Skrzpczynski, Z., Allawi, H.T., Wayland, S.R., Takova, T. and Neri, B.P.  
TITLE Charge tags and the separation of nucleic acid molecules  
JOURNAL Patent: US 6780982-A 72 24-AUG-2004;  
FEATURES Location/Qualifiers  
source 1..647  
/organism="unknown"  
/mol\_type="unassigned RNA"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 647;  
Best Local Similarity 99.0%; Pred. No. 3.8e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 60  
|||||  
Db 326 TGCAGCTAACTATTTCCCTAGCTTCCCGACAGACCTGTGTTTATTATTATAATG 385  
|||||

QY 61 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 120  
|||||

Db 386 AATTTGTTGTTGATGTAACAATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 445  
|||||

QY 121 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 180  
|||||

Db 446 TTGATGTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGA 505  
|||||

QY 181 AATTGCTTTTCCCTCTGTAC 200  
|||||

Db 506 AATTGCTTTTCCCTCTGTGAAC 525  
|||||

RESULT 13  
LOCUS AX698741 647 bp RNA linear PAT 02-APR-2003  
DEFINITION Sequence 72 from Patent WO2003030.  
ACCESSION AX698741  
VERSION AX698741.1 GI:29499529  
KEYWORDS synthetic construct  
SOURCE other sequences; artificial sequences.  
ORGANISM 1  
REFERENCE 1  
AUTHORS Lyanichev, V., Skrzpczynski, Z., Allawi, H.T., Wayland, S.R., Takova, T. and Neri, B.P.  
TITLE Charge tags and separation of nucleic acid molecules  
JOURNAL Patent: WO 02063030-A 72 15-AUG-2002;  
THIRD WAVE TECHNOLOGIES, INC. (US)  
FEATURES Location/Qualifiers  
source 1..647  
/organism="synthetic construct"

Tue Aug 9 17:01:25 2005

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 647;  
Best Local Similarity 99.0%; Pred. No. 3.7e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTATTATAATG 60  
326 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTGTTTATTATTATAATG 385

DB

QY 61 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 120  
386 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 445

DB

QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 180  
446 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 505

DB

QY 181 AATTGCTTTTCCCTCTGTAC 200  
506 AATTGCTTTTCCCTCTGTAC 525

DB

RESULT 14  
AR337874 725 bp DNA linear PAT 17-AUG-2003  
LOCUS  
DEFINITION Sequence 3 from patent US 6569418.  
ACCESSION AR337874  
VERSION AR337874.1 GI:33724486  
KEYWORDS  
SOURCE  
ORGANISM  
Unkown.  
Unclassified.  
REFERENCE 1 (bases 1 to 725)  
AUTHORS Garzino-Demo,A. and Devico,A.L.  
TITLE Immuno-modulating effects of chemokines in DNA vaccination  
JOURNAL Patent: US 6569418-A 3 27-MAY-2003;  
FEATURES  
source  
1..725  
/organism="unkown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 725;  
Best Local Similarity 99.0%; Pred. No. 3.7e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTATTATAATG 60  
379 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTGTTTATTATTATAATG 438

DB

QY 61 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 120  
439 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 498

DB

QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 180  
499 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 558

DB

QY 181 AATTGCTTTTCCCTCTGTAC 200  
559 AATTGCTTTTCCCTCTGTAC 578

DB

SOURCE Unknown.  
ORGANISM Unkown.  
REFERENCE 1 (bases 1 to 725)  
AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.  
TITLE Compositions for the detection of blood cell and immunological response gene expression  
JOURNAL Patent: US 6607879-A 1344 19-AUG-2003;  
FEATURES  
source  
1..725  
/organism="unkown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 98.4%; Score 196.8; DB 6; Length 725;  
Best Local Similarity 99.0%; Pred. No. 3.7e-26;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTATTATAATG 60  
379 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTGTTTATTATTATAATG 438

DB

QY 61 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 120  
439 AATTTTGTGTTGATGTAACATTAATGCTTAAGTAATGTTTATTATTATAATG 498

DB

QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 180  
499 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 558

DB

QY 181 AATTGCTTTTCCCTCTGTAC 200  
559 AATTGCTTTTCCCTCTGTAC 578

DB

Search completed: August 4, 2005, 11:16:39  
Job time : 807.887 secs

RESULT 15  
AR380799 725 bp DNA linear PAT 18-DEC-2003  
LOCUS  
DEFINITION Sequence 1344 from patent US 6607879.  
ACCESSION AR380799  
VERSION AR380799.1 GI:40088433  
KEYWORDS

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 132.551 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-2

Perfect score: 200

Sequence: 1 tgcagtaactattttccc.....aattgtttctctgttac 200

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	200	100.0	200	6	ADH13939 Human mon
2	198.4	99.2	489	9	ACH19135 Human adu
3	198.4	99.2	772	13	ADS16434 Human cyt
C 4	198.4	99.2	1005	10	ADI02670 Human CDN
C 5	198.4	99.2	1472	10	ADE25637 Human CDN
6	198.4	99.2	11793	12	ADO03803 Human Ccl
7	198	99.0	950	8	ACF64400 Human MCP
8	197.6	98.8	2775	6	ABK47979 Human sma
C 9	196.8	98.4	360	4	AAI15568 Probe #55
C 10	196.8	98.4	360	4	ABA57500 Human foe
C 11	196.8	98.4	360	4	AAI17053 Probe #57
C 12	196.8	98.4	360	4	ABA26988 Probe #54
C 13	196.8	98.4	360	4	AAK31146 Human bon
C 14	196.8	98.4	360	4	AAK05543 Human bra
C 15	196.8	98.4	360	4	ABS30826 Human liv
C 16	196.8	98.4	360	6	ABS05898 Human gen
17	196.8	98.4	508	13	ADR52943 Drug ther
18	196.8	98.4	647	6	ABS68800 Human mon
19	196.8	98.4	725	2	AAQ85370 Chemoatr
20	196.8	98.4	725	2	AAx80631 Monocyte

21	196.8	98.4	725	10	AAD58817 Human mon
22	196.8	98.4	725	10	ADD18494 Human pro
23	196.8	98.4	725	10	AdE84839 Farnesyl
24	196.8	98.4	725	11	ADI32018 Human CDN
25	196.8	98.4	725	13	ADR05375 Human mon
26	196.8	98.4	738	1	AAN91337 DNA which
27	196.8	98.4	739	2	AAV10341 cDNA enco
28	196.8	98.4	739	6	ABV78071 Hypoxia-r
29	196.8	98.4	739	8	ABZ34738 Coding se
30	196.8	98.4	739	12	AdF45446 Human vas
31	196.8	98.4	739	12	Adn04745 Antipsoi
32	196.8	98.4	739	13	ADR24988 Breast ca
33	196.8	98.4	739	13	ADP24778 PRO poly
34	196.8	98.4	741	3	AAA34899 Human ade
35	196.8	98.4	741	3	Aaf21021 Human low
36	196.8	98.4	741	10	ABZ96715 Human nuc
37	196.8	98.4	741	11	ADI31839 Human CDN
38	196.8	98.4	741	11	ABD20564 Human pul
39	196.8	98.4	743	12	ADO03801 Human Ccl
40	196.8	98.4	756	10	ADI02669 Human CDN
41	196.8	98.4	757	10	ADD14996 Human mon
42	196.8	98.4	757	11	ADN95635 Human BEC
43	196.8	98.4	757	12	ADO03800 Human Ccl
44	196.8	98.4	757	12	ADP75913 Human SCY
45	196.8	98.4	757	12	ADQ76211 Chemokine

## ALIGNMENTS

## RESULT 1

ADH13939

ID ADH13939 standard; DNA; 200 BP.

AC ADH13939;

DT 11-MAR-2004 (first entry)

DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:2.

ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
cervical carcinoma.

OS Homo sapiens.

XX EP1170372-A1.

XX 09-JAN-2002.

PF 06-JUL-2000; 2000EP-00114560.

XX 06-JUL-2000; 2000EP-00114560.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

PI Roehl F, Soto U, Coy J, Finzer P, Delius H, Poustka A;

PI Zur Hausen H, Patzelt A;

DR WPI; 2002-165895/22.

PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
useful in gene therapy, for treating atherosclerosis and cancer.

XX Disclosure; SEQ ID NO 2; 30pp; English.

PS The invention relates to a novel nucleic acid molecule (I) comprising a  
sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
having the biological activity of MCP-1. A protein encoded by a nucleic  
acid of the invention has cytostatic, and antiarteriosclerotic activity.  
CC A nucleic acid of the invention may have a use in gene therapy. A  
compound of the invention is useful in the preparation of a medicament  
for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

Tue Aug 9 17:01:25 2005

CC pharmaceutical composition of the invention is useful for the treatment  
CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
CC atherosclerosis or cancer. The present sequence is used in the  
CC exemplification of the invention.

XX Sequence 200 BP; 48 A; 29 C; 31 G; 92 T; 0 U; 0 Other;  
SQ

Query Match 100.0%; Score 200; DB 6; Length 200;  
Best Local Similarity 100.0%; Pred. No. 1.5e-30;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTCTTTTATTATTATAAG 60  
DB 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTCTTTTATTATTATAAG 60

QY 61 AATTTTGTGTTGATGTGAACAATTATGCTTAAGTAATGTTAATTTTAAGTTA 120  
DB 61 AATTTTGTGTTGATGTGAACAATTATGCTTAAGTAATGTTAATTTTAAGTTA 120

QY 121 TTGATGTTTAAAGTTATCTTTCATGCTAGTCTTTTATAGATACAGAGACTTGGGGA 180  
DB 121 TTGATGTTTAAAGTTATCTTTCATGCTAGTCTTTTATAGATACAGAGACTTGGGGA 180

QY 181 AATTCCTTTTCCCTCTGTAC 200  
DB 181 AATTCCTTTTCCCTCTGTAC 200

RESULT 2  
ACH19135  
ID ACH19135 standard; cDNA; 489 BP.  
XX  
AC ACH19135;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human adult lung cDNA #138.  
XX  
KW Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2003073623-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 30-JUL-2001; 2001US-00918995.  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
FT as hybridization probes, as oligomers for PCR, for chromosome and gene  
FT mapping, in the recombinant production of protein, or in generating  
FT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 6347; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SHH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences

CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The present polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
SQ Sequence 489 BP; 147 A; 95 C; 75 G; 165 T; 0 U; 7 Other;  
Query Match 99.2%; Score 198.4; DB 9; Length 489;  
Best Local Similarity 99.5%; Pred. No. 3.1e-30;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTGTTTATTATTATAAG 60  
DB 139 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTGTTTATTATTATAAG 198

QY 61 AATTTTGTGTTGATGTGAACAATTATGCTTAAGTAATGTTAATTTTAAGTTA 120  
DB 199 AATTTTGTGTTGATGTGAACAATTATGCTTAAGTAATGTTAATTTTAAGTTA 258

QY 121 TTGATGTTTAAAGTTATCTTTCATGCTAGTCTTTTATAGATACAGAGACTTGGGGA 180  
DB 259 TTGATGTTTAAAGTTATCTTTCATGCTAGTCTTTTATAGATACAGAGACTTGGGGA 318

QY 181 AATTCCTTTTCCCTCTGTAC 200  
DB 319 AATTCCTTTTCCCTCTGTAC 338

RESULT 3  
ADS16434  
ID ADS16434 standard; cDNA; 772 BP.  
XX  
AC ADS16434;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human cytokine, JE cDNA.  
XX  
KW Human; cytokine; JE; immune responsiveness; wound healing; tissue repair;  
KW bacterial infection; viral infection; anaemia; B cell deficiency;  
KW T cell deficiency; vulnerability; antibacterial; virucide; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 73..372  
XX /\*tag= c  
XX /product= "JE cytokine"  
XX  
XX sig\_peptide 73..159  
XX /\*tag= b  
XX sig\_peptide 73..141  
XX /\*tag= a  
XX mat\_peptide 142..369  
XX /\*tag= d  
XX /product= "Mature cytokine factor, JE"  
XX mat\_peptide 160..369  
XX /\*tag= e  
XX /product= "Mature cytokine factor, JE"  
XX  
XX US6787645-B1.  
XX  
XX 07-SEP-2004.  
XX  
XX 09-MAY-1995; 95US-00437306.  
XX

```

PR 12-MAY-1989; 89US-00351008.
PR 16-MAY-1991; 91US-00701515.
PR 12-JAN-1993; 93US-00003136.
PR 13-APR-1994; 94US-00228931.
XX
PA (DAND ) DANA FARBEN CANCER INST INC.
PA (GEM ) GENETICS INST INC.
XX
PI Rollins BJ, Stiles CD, Wong GG;
XX
XX WPI; 2004-632933/61.
DR P-PSDB; ADS16435.
XX
XX New isolated human JE cytokine DNA and polypeptide, useful for treating
PT bacterial and viral infections, anemia, or B cell or T cell deficiencies,
PT or in wound healing and related tissue repair.
XX
XX Claim 1; SEQ ID NO 1; 8pp; English.
XX
XX The invention relates to a nucleic acid sequence encoding human cytokine,
CC JE. JE sequence is useful in pharmaceutical preparations for stimulating
CC and/or enhancing immune responsiveness, wound healing and related tissue
CC repair. It is also useful for treating bacterial and viral infections,
CC anaemia or B/T cell deficiencies. The present sequence is human JE cDNA.
XX
XX Sequence 772 BP; 239 A; 175 C; 134 G; 224 T; 0 U; 0 Other;
SQ
Query Match 99.2%; Score 198.4; DB 13; Length 772;
Best Local Similarity 99.5%; Pred. No. 3.1e-30;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTTTATTTATTAATG 60
DB TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTTTATTTATTAATG 457
QY 61 AATTTTGTGTTGATGTGAACATTATGCTTAACTTAAGTAAATCTTATTAAGTTA 120
DB AATTTTGTGTTGATGTGAACATTATGCTTAACTTAAGTAAATCTTATTAAGTTA 517
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
DB TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 577
QY 181 AATTGCTTTTCCCTCTGTAC 200
DB AATTGCTTTTCCCTCTGTGAAC 597
RESULT 4
AD102670/c
ID AD102670 standard; cDNA; 1005 BP.
XX
XX AD102670;
XX
XX 22-APR-2004 (first entry)
XX
XX Human cDNA differentially expressed in the vascular endothelium #211.
XX
XX ss; vascular endothelium; vascular disorder; atherosclerosis;
XX haemangioma; haemangioendothelioma; oedema; diabetic retinopathy; wart;
XX pyogenic granuloma; Kaposi's sarcoma; scar keloid; allergic oedema;
XX neoplasm; psoriasis; ulcer; follicular cyst; endometriosis;
XX peritoneal sclerosis; obesity; human.
XX
XX Homo sapiens.
XX
XX US2003166903-A1.
XX
XX 04-SEP-2003.
XX
XX 25-APR-2002; 2002US-00133013.
XX
XX 27-APR-2001; 2001US-0287067P.
XX

```

```

XX (ASTR/) ASTROMOFF A.
PA (BAND/) BANDMAN O.
PA (COCK/) COCKS B G.
XX
XX Astromoff A, Bandman O, Cocks BG;
XX
XX WPI; 2003-898115/82.
XX
XX New combination comprising cDNAs that are differentially expressed in
PT vascular endothelium, useful for preparing a composition for diagnosing
PT or treating vascular disorder, e.g., atherosclerosis, psoriasis, ulcers
PT or obesity.
XX
XX Claim 1; SEQ ID NO 211; 28pp; English.
XX
XX The invention relates to a new combination which comprises cDNAs or their
CC complements not given in the specification that are differentially
CC expressed in vascular endothelium. The combination of cDNAs is useful for
CC preparing a composition for diagnosing or treating vascular disorder,
CC comprising atherosclerosis, haemangioma, haemangioendothelioma, oedema,
CC diabetic retinopathy, warts, pyogenic granulomas, Kaposi's sarcoma, scar
CC keloids, allergic oedema, neoplasms, psoriasis, ulcers, follicular cysts,
CC endometriosis, peritoneal sclerosis or obesity. The present sequence
CC represents a cDNA differentially expressed in the vascular endothelium.
XX
XX Sequence 1005 BP; 318 A; 183 C; 178 G; 326 T; 0 U; 0 Other;
SQ
Query Match 99.2%; Score 198.4; DB 10; Length 1005;
Best Local Similarity 99.5%; Pred. No. 3.1e-30;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTTTATTTATTAATG 60
DB TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTTGTTTATTTATTAATG 288
QY 61 AATTTTGTGTTGATGTGAACATTATGCTTAACTTAAGTAAATCTTATTAAGTTA 120
DB AATTTTGTGTTGATGTGAACATTATGCTTAACTTAAGTAAATCTTATTAAGTTA 228
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
DB TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 168
QY 181 AATTGCTTTTCCCTCTGTAC 200
DB AATTGCTTTTCCCTCTGTGAAC 148
RESULT 5
ADE25637/c
ID ADE25637 standard; cDNA; 1472 BP.
XX
XX ADE25637;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human cDNA differentially expressed in foam cells #41.
XX
XX Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;
XX cardiovascular disease; atherosclerosis.
XX
XX Homo sapiens.
XX
XX US2003194721-A1.
XX
XX 16-OCT-2003.
XX
XX 18-SEP-2002; 2002US-00247671.
XX
XX 19-SEP-2001; 2001US-0323784P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX

```

**Tue Aug 9 17:01:25 2005**

XX Mikita T, Shiffman D, Porter JG, Kaser MR;  
PI WPI; 2003-875398/81.  
XX  
XX Combination containing several polynucleotides that are differentially  
XX expressed in foam cells and complements of the polynucleotides, useful  
XX for diagnosing cardiovascular disease or atherosclerosis.  
XX  
XX Claim 1; SEQ ID NO 41; 37pp; English.  
PS  
XX The invention relates to a combination comprising several polynucleotides  
XX having any one of 127 sequences (S1) such as the sequence of human  
CC calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4  
CC hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit  
CC mRNA, etc., and their complements. The cDNAs are differentially expressed  
CC in LPS (lipopolysaccharide)-treated foam cells. Also included are  
CC obtaining an extended or full length gene from a library of nucleic acid  
CC sequences, an expression vector containing the nucleic acids, a host cell  
CC containing the vector, a purified polypeptide appearing as AGE25750 and  
CC AGE23751, producing a protein by culturing the host cell, and a  
CC composition comprising a purified antibody that specifically binds to the  
CC polynucleotides in a sample. The sample is from a subject with  
CC atherosclerosis and comparison with a standard defines early, mid or late  
CC stages of the disorder. The foam cell-expressed nucleic acids are useful  
CC for high throughput screening of a library of molecules or compounds to  
CC identify a ligand which binds a polynucleotide. The library is chosen  
CC from DNA molecules, peptides, proteins and RNA molecules. The protein is  
CC useful for a high throughput screening of library of molecules or  
CC compounds to identify at least one ligand which specifically binds a  
CC protein, for purifying a ligand from a sample for making an antibody. The  
CC foam cell-expressed nucleic acids are useful for diagnosing  
CC cardiovascular disorder. The foam cell-expressed nucleic acids are useful  
CC as elements on a microarray which can be used for detecting related  
CC polynucleotide in a sample, diagnosing cardiovascular disease,  
CC atherosclerosis. The present sequence represents a cDNA whose expression  
CC is upregulated in LPS treated foam cells.  
XX  
XX Sequence 1472 BP; 462 A; 303 C; 268 G; 439 T; 0 U; 0 Other;  
SQ

Query Match	99.2%;	Score 198.4;	DB 10;	Length 14/27;
Best Local Similarity	99.5%;	Pred. NO. 3e-30;		
Matches 199;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	TGGAGCTAACTTATTTTCCCTAGCTTTCCCGACACACCTCTGTTTATTTTATTAATG	60	
Db	347	TGCGACTAACCTTATTTTCCCTAGCTTTCCCGACACACCTGTTTATTTTATTAATG	288	
QY	61	AATTTTGTTGTCGATGTGAACCAATTAATGCTTAAAGTAATGTTAATCTTATTTAAGTTA	120	
Db	287	AATTTTGTTGTTGATGTGAACCAATTAATGCTTAAAGTAATGTTAATCTTATTTAAGTTA	228	
QY	121	TTGATGTTTTAAGTTTATCTTTGATGGTACTAGTGTGTTTTTACATACAGAGACTTGGGGA	180	
Db	227	TTGATGTTTTAAGTTTATCTTTTATCTTTTATGTTACTAGTGTGTTTTTACATACAGAGACTTGGGGA	168	
QY	181	AATTGCTTTTTCCTCTGTAC	200	
Db	167	AATTGCTTTTTCCTCTTGAAC	148	

RESULT 6	
ADO03803	
ID	ADO03803 standard; DNA; 11793 BP.
XX	
XX	
AC	ADO03803;
XX	
DT	12-AUG-2004 (first entry)
XX	
XX	
DE	Human Cc12 gene and enhancer regi
XX	
XX	

human; ds; animal model; age-related macular degeneration; and  
gene knockout; Cx12-deficient; Cx2-deficient; drusen;  
lipofuscin accumulation; Bruch's membrane; retinal degeneration;  
choroidal neovascularisation; ophthalmological; gene therapy.

Homo sapiens.

WO2004041160-A2.

21-MAY-2004.

16-OCT-2003; 2003WO-US032933.

30-OCT-2002; 2002US-0422096P.

(KENT ) UNIV KENTUCKY RES FOUND.

Ambati J;

WFI; 2004-400512/37.

Testing candidate drug for treating age-related macular degeneration, by  
administering drug to Cx12-deficient, Cx2-deficient knockout mouse, and  
analyzing development or regression of drusen and/or lipofuscin  
accumulation in eye.

Disclosure; SEQ ID NO 4; 64pp; English.

This invention relates to a novel methods and animal models for testing  
candidate drugs that can be used for the treatment or prevention of age-  
related macular degeneration (AMD). Specifically, it refers to  
administering a candidate drug to gene knockout mice, in particular Cx12-  
deficient, Cx2-deficient and/ or Cx12-deficient/Cx2-deficient dual  
knockout mouse. The present invention describes analysing the knockout  
mouse eye for development or regression of drusen and/ or lipofuscin  
accumulation, as well as for the effect of a candidate drug on Bruch's  
membrane, retinal degeneration and/ or choroidal neovascularisation.  
Accordingly, such compositions exhibit ophthalmological activities and  
can be used for gene therapy purposes. This polynucleotide sequence is  
the human Cx12 gene and enhancer region DNA of the invention.

Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;

Query Match 99.2%; Score 198.4; DB 12; Length 11793;  
Best Local Similarity 99.5%; Pred. No. 3e-30;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACACCTTGTTTATTATTATTAATG 60  
9099 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACACCTTGTTTATTATTATTAATG 9158

61 AATTTGTTTGTGATGTGAACACATTATGCCTTAAGTAATGTTATCTTATTAAAGTTA 120  
9159 AATTTGTTTGTGATGTGAACACATTATGCCTTAAGTAATGTTATCTTATTAAAGTTA 9218

121 TTGATGTTTAAAGTTATCTTTCATGCTAGTACGTAGTGTGTTTTTAGATACAGAGACTCGGGA 180  
9219 TTGATGTTTAAAGTTATCTTTCATGCTAGTACGTAGTGTGTTTTTAGATACAGAGACTCGGGA 9278

181 AATTCGCTTTCCCTCTGTAC 200

9279 AATTCGCTTTCCCTCTGTAC 9298

RESULT 7  
ACF64400  
ID ACF64400 standard; DNA; 950 BP.  
AC ACF64400;  
XX  
XX 13-OCT-2003 (first entry)  
DT  
DE Human MCP1 nucleotide sequence >MCP1\_03.

```
XX Human; detection; computer-readable storage medium; polymorphic site;
KW signal carrying data; data processing system; multiple sclerosis; gene;
KW ds.
XX
XX Homo sapiens.
OS
XX WO2003014319-A2.
XX
XX 20-FEB-2003.
XX
XX 07-AUG-2002; 2002WO-US025268.
XX
XX 07-AUG-2001; 2001US-0310741P.
XX
XX 24-SEP-2001; 2001US-0324790P.
XX
XX (DNAS-) DNA SCI INC.
XX
XX Jones HB, Xu H, White R, Rienhoff HY, Jin W, Natsoulis G;
XX WPI; 2003-268196/26.
XX
XX New polynucleotide, useful for detecting loci associated with multiple
XX sclerosis.
XX
XX Disclosure; Page 75; 93pp; English.
XX
XX The present invention describes an isolated polynucleotide (PN)
XX comprising: (a) a sequence comprising at least 15 contiguous nucleotides
XX of a sequence comprising variant sequences (A) from Table 4 given in the
XX specification; or (b) a sequence that is complementary to (A). Also
XX described: (1) an array of (PN)s comprising two or more of the isolated
XX (PN)s; (2) detecting a (PN) in an individual; (3) a computer-readable
XX storage medium where each record has a field identifying a base
XX occupying a (PN) site and a location of the polymorphic site; and (4) a
XX signal carrying data for access by an application program having executed
XX on a data processing system. The (PN) can be used for detecting loci
XX associated with multiple sclerosis. ACF64025 to ACF64424 represent
XX sequences used in the exemplification of the present invention
XX
XX Sequence 950 BP; 255 A; 220 C; 184 G; 290 T; 0 U; 1 Other;
XX
XX Query Match 99.0%; Score 198; DB 8; Length 950;
XX Best Local Similarity 99.0%; Pred. No. 3.7e-30;
XX Matches 198; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACACCTTGTATTTATTTATTAATG 60
XX 357 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACACCTTGTATTTATTTATTAATG 416
XX
XX 61 AATTTTGTGTTGATGTGAACATTATGCTTTAAGTAATGTTAATTTCTTTAAGTTA 120
XX 417 AATTTTGTGTTGATGTGAACATTATGCTTTAAGTAATGTTAATTTCTTTAAGTTA 476
XX
XX 121 TTGATGTTTTAAGTTATCTTTTCATGGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 180
XX 477 TTGATGTTTTAAGTTATCTTTTCATGGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 536
XX
XX 181 AATTCCTTTTCTCTGTGAC 200
XX 537 AATTCCTTTTCTCTGTGAC 556
XX
XX RESULT 8
XX ABK47979
XX ID ABK47979 standard; DNA; 2775 BP.
XX
XX AC ABK47979;
XX
XX XX
XX 02-JUL-2002 (first entry)
XX
XX Human small inducible cytokine A2 (SCYA2) genomic DNA.
XX
```

```
KW Human; small inducible cytokine A2; SCYA2; gene; ds; haplotype pair;
KW haplotyping; atherosclerosis; antiarteriosclerotic; gene therapy;
KW single nucleotide polymorphism; genotyping; drug screening;
KW chromosome 17q11.2-q21.1.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 598..2080
XX FT /*tag= a
XX FT /product= "Human SCYA2"
XX
XX WO200218413-A2.
XX
XX 07-MAR-2002.
XX
XX 28-AUG-2001; 2001WO-US026899.
XX
XX 28-AUG-2000; 2000US-0228496P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Finkel K, Koshy B, Kumar AM, Lee HH;
XX WPI; 2002-339655/37.
XX P-PSDB; AAU77179.
XX
XX New genetic variants having polymorphisms in the small inducible cytokine
XX A1 (SCYA2) gene, useful for studying the function of SCYA2, and for
XX treating disorders affected by expression or function of the SCYA2
XX isogene.
XX
XX Claim 1; Fig 1; 58pp; English.
XX
XX The invention relates to single nucleotide polymorphisms in the gene
XX encoding human small inducible cytokine A2 (SCYA2) polypeptide. A method
XX for haplotyping the SCYA2 gene in an individual comprises identifying the
XX nucleotide at one or more polymorphic sites and determining whether one
XX of the copies of the gene is defined by one of the SCYA2 haplotypes given
XX in the specification or whether both copies are defined by a haplotype
XX pair. This method is useful in genotyping, whereby all possible haplotype
XX pairs can be assigned to specific genotypes. An association between a
XX trait and a haplotype or haplotype pair of the SCYA2 gene can be
XX identified by comparing the frequency of the haplotype or haplotype pair
XX in a population exhibiting the trait with the frequency of the haplotype
XX or haplotype pair in a reference population, where a higher haplotype
XX frequency in the trait population indicates the trait is associated with
XX the haplotype or haplotype pair. SCYA2 and its corresponding DNA are used
XX for studying the expression and function of SCYA2, and in screening for
XX candidate drugs to treat diseases related to SCYA2 activity, such as
XX atherosclerosis. This sequence represents genomic DNA which encodes the
XX human SCYA2 polypeptide
XX
XX Sequence 2775 BP; 699 A; 723 C; 565 G; 778 T; 0 U; 10 Other;
XX
XX Query Match 98.8%; Score 197.6; DB 6; Length 2775;
XX Best Local Similarity 98.5%; Pred. No. 4.3e-30;
XX Matches 197; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACACCTTGTGTTTTTATTTATTAATG 60
XX 2106 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACACCTTGTGTTTTTATTTATTAATG 2165
XX
XX 61 AATTTTGTGTTGATGTGAACATTATGCTTTAAGTAATGTTAATTTCTTTAAGTTA 120
XX
XX 2166 AATTTTGTGTTGATGTGAACATTATGCTTTAAGTAATGTTAATTTCTTTAAGTTA 2225
XX
XX 121 TTGATGTTTTAAGTTATCTTTTCATGGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 180
XX
XX 2226 TTGATGTTTTAAGTTATCTTTTCATGGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 2285
XX
XX 181 AATTCCTTTTCTCTGTGAC 200
XX
XX ||||| :||
```

Tue Aug 9 17:01:25 2005

us-09-899-276c-2.rng

```

Db      2286 AATTGCTTTTCTCTTGAC 2305
RESULT 9
AA115568/C
ID   AA115568 standard; DNA; 360 BP.
XX
AC   AA115568;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Probe #5501 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
XX   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000670.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX   26-MAY-2000; 2000US-0207456P.
XX   30-JUN-2000; 2000US-00608408.
XX   03-AUG-2000; 2000US-00632366.
XX   21-SEP-2000; 2000US-0234687P.
XX   27-SEP-2000; 2000US-0236359P.
XX   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX   WPI; 2001-488901/53.
XX
DR   Human genome-derived single exon nucleic acid probes useful for analyzing
XX   gene expression in human cervical epithelial cells.
XX
PS   Claim 25; SEQ ID NO 5501; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging of
CC   diseases of the cervix, notably cervical cancer. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match      98.4%; Score 196.8; DB 4; Length 360;
Best Local Similarity 99.0%; Pred. No. 6.4e-30;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1   TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 60
Db      227   TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTGTTTATTTTATTAATG 168
QY      61   AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTTTATTTTATTAAGTTA 120
Db      167   AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTTTATTTTATTAAGTTA 108
QY      121   TTGATGTTTAAAGTTTATCTTTTCATGTTACTAGTGTGTTTTTATAGATACAGAGACTTGGGA 180
Db      107   TTGATGTTTAAAGTTTATCTTTTCATGTTACTAGTGTGTTTTTATAGATACAGAGACTTGGGA 48
QY      181   AATTGCTTTTCTCTTGAC 200
Db      47   AATTGCTTTTCTCTTGAC 28

Db      119 AATTGCTTTTCTCTTGAC 28
RESULT 10
ABA57500/C
ID   ABA57500 standard; DNA; 360 BP.
XX
AC   ABA57500;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #5805.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX   26-MAY-2000; 2000US-0207456P.
XX   30-JUN-2000; 2000US-00608408.
XX   03-AUG-2000; 2000US-00632366.
XX   21-SEP-2000; 2000US-0234687P.
XX   27-SEP-2000; 2000US-0236359P.
XX   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX   WPI; 2001-483447/52.
XX
DR   Human genome-derived single exon nucleic acid probes useful for analyzing
XX   gene expression in human fetal liver.
XX
PS   Claim 1; SEQ ID NO 5805; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human fetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match      98.4%; Score 196.8; DB 4; Length 360;
Best Local Similarity 99.0%; Pred. No. 6.4e-30;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1   TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 60
Db      227   TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTGTTTATTTTATTAATG 168
QY      61   AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTTTATTTTATTAAGTTA 120
Db      167   AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTTTATTTTATTAAGTTA 108
QY      121   TTGATGTTTAAAGTTTATCTTTTCATGTTACTAGTGTGTTTTTATAGATACAGAGACTTGGGA 180
Db      107   TTGATGTTTAAAGTTTATCTTTTCATGTTACTAGTGTGTTTTTATAGATACAGAGACTTGGGA 48
QY      181   AATTGCTTTTCTCTTGAC 200
Db      47   AATTGCTTTTCTCTTGAC 28

```

```
RESULT 11
AAI37053/c
ID AAI37053 standard; DNA; 360 BP.
XX
AC AAI37053;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #5739 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 5739; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;
Query Match 98.4%; Score 196.8; DB 4; Length 360;
Best Local Similarity 99.0%; Pred. No. 6.4e-30;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCCAGACACCTTGTTTATTTATTAATG 60
DB 227 TGCAGCTAACTTATTTCCCTAGCTTTCCCCAGACACCTTGTTTATTTATTAATG 168
QY 61 AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTAATCTTATTTAAGTTA 120
DB 167 AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTAATCTTATTTAAGTTA 108
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTATAGATACAGAGACTTGGGA 180
DB 107 TTGATGTTTAAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTATAGATACAGAGACTTGGGA 48
QY 181 AATTCCTTTTCCCTCTGTAC 200
DB 47 AATTCCTTTTCCCTCTGTAC 28
```

```
ID ABA26988 standard; DNA; 360 BP.
XX
AC ABA26988;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #5454 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 1; SEQ ID NO 5454; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;
Query Match 98.4%; Score 196.8; DB 4; Length 360;
Best Local Similarity 99.0%; Pred. No. 6.4e-30;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCCAGACACCTTGTTTATTTATTAATG 60
DB 227 TGCAGCTAACTTATTTCCCTAGCTTTCCCCAGACACCTTGTTTATTTATTAATG 168
QY 61 AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTAATCTTATTTAAGTTA 120
DB 167 AATTTGTTTGTGATGTGAACATTATGCTTAAGTAATGTAATCTTATTTAAGTTA 108
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTATAGATACAGAGACTTGGGA 180
DB 107 TTGATGTTTAAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTATAGATACAGAGACTTGGGA 48
QY 181 AATTCCTTTTCCCTCTGTAC 200
DB 47 AATTCCTTTTCCCTCTGTAC 28
```

Tue Aug 9 17:01:25 2005

AAK05543 standard; DNA; 360 BP.

ID

XX

AC

AAK05543;

XX

05-NOV-2001 (first entry)

DT

XX

DE

Human brain expressed single exon probe SEQ ID NO: 5534.

XX

Human; brain expressed exon; gene expression analysis; probe; microarray; Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer; ss.

XX

OS

Homo sapiens.

XX

WO200157275-A2.

PN

XX

XX

09-AUG-2001.

PD

XX

XX

30-JAN-2001; 2001WO-US000667.

PF

XX

XX

04-FEB-2000; 2000US-0180312P.

PR

26-MAY-2000; 2000US-0207456P.

PR

30-JUN-2000; 2000US-00608408.

PR

03-AUG-2000; 2000US-00632366.

PR

21-SEP-2000; 2000US-0234687P.

PR

27-SEP-2000; 2000US-0236359P.

PR

04-OCT-2000; 2000GB-00024263.

XX

(MOLE-) MOLECULAR DYNAMICS INC.

PA

Penn SG, Hanzel DK, Chen W, Rank DR;

XX

WPI; 2001-483446/52.

XX

Single exon nucleic acid probes for analyzing gene expression in human brains.

XX

Example 4; SEQ ID NO 5534; 650pp + Sequence Listing; English.

XX

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention

XX

Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

SQ

Query Match 98.4%; Score 196.8; DB 4; Length 360;  
Best Local Similarity 99.0%; Pred. No. 6.4e-30;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy

1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 60

Db

227 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 168

Qy

61 AATTTGTTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTAAGTTA 120

Db

167 AATTTGTTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTAAGTTA 108

Qy

121 TTGATGTTTAAAGTTATCTTTCATGTTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 180

Db

107 TTGATGTTTAAAGTTATCTTTCATGTTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 48

Qy

181 AATTCCTTTTCCCTCTGTAC 200

Db

47 AATTCCTTTTCCCTCTGTAC 28

RESULT 15

ABS30826/c

ID

ABS30826 standard; DNA; 360 BP.

AAK31146/c

AAK31146 standard; DNA; 360 BP.

RESULT 13

ID

AAK31146;

XX

06-NOV-2001 (first entry)

DT

XX

DE

Human bone marrow expressed single exon probe SEQ ID NO: 5703.

XX

Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX

OS

Homo sapiens.

XX

WO200157276-A2.

PN

XX

XX

09-AUG-2001.

PD

XX

XX

30-JAN-2001; 2001WO-US000668.

PF

XX

XX

04-FEB-2000; 2000US-0180312P.

PR

26-MAY-2000; 2000US-0207456P.

PR

30-JUN-2000; 2000US-00608408.

PR

03-AUG-2000; 2000US-00632366.

PR

21-SEP-2000; 2000US-0234687P.

PR

27-SEP-2000; 2000US-0236359P.

PR

04-OCT-2000; 2000GB-00024263.

XX

(MOLE-) MOLECULAR DYNAMICS INC.

PA

Penn SG, Hanzel DK, Chen W, Rank DR;

XX

WPI; 2001-488900/53.

XX

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow.

XX

Example 4; SEQ ID NO 5703; 658pp + Sequence Listing; English.

XX

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is one of the probes of the invention

XX

Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

SQ

Query Match 98.4%; Score 196.8; DB 4; Length 360;  
Best Local Similarity 99.0%; Pred. No. 6.4e-30;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy

1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 60

Db

227 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATG 168

Qy

61 AATTTGTTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTAAGTTA 120

Db

167 AATTTGTTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTAAGTTA 108

Qy

121 TTGATGTTTAAAGTTATCTTTCATGTTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 180

Db

107 TTGATGTTTAAAGTTATCTTTCATGTTACTAGTGTGTTTATAGATACAGAGACTTGGGGA 48

Qy

181 AATTCCTTTTCCCTCTGTAC 200

Db

47 AATTCCTTTTCCCTCTGTAC 28

RESULT 14

AAK05543/c

```
XX ABS30826;
XX AC
XX DT
XX 25-FEB-2003 (first entry)
XX DE
XX Human liver single exon probe, SEQ ID No 5816.
XX KW
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157273-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000664.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX PA
XX (MOLS-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WIPI; 2001-488898/53.
XX DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX PS
XX Claim 1; SEQ ID NO 5816; 658pp; English.
XX CC
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS51005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match 98.4%; Score 196.8; DB 4; Length 360;
Best Local Similarity 99.0%; Pred. No. 6.4e-30;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTTTATTTTATTAATG 60
Db |||||||
227 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTTTATTTTATTAATG 168
QY 61 AATTTGTTGTTGATGAAACATTTATGCTTAACTTAATCTTATTTAAGTTA 120
Db |||||||
167 AATTTGTTGTTGATGTTGAAACATTTATGCTTAACTTAATCTTATTTAAGTTA 108
QY 121 TTGATGTTTAACTTATCTTTTCATGCTAGTGTGTTTATGATACAGAGACTTGGGGA 180
Db |||||||
107 TTGATGTTTAACTTATCTTTTCATGCTAGTGTGTTTATGATACAGAGACTTGGGGA 48
QY 181 AATTGCTTTTCCTCTTTGTAC 200
Db |||||||
```

Db 47 AATTGCTTTTCCTCTTTGAAC 28

Search completed: August 4, 2005, 08:31:05  
Job time : 139.551 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 943.968 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-2

Perfect score: 200

Sequence: 1 tgcagtaactattttccc.....aattgtttctcttctgtac 200

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gssi:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	198.4	99.2	424	5	BU729482	UI-B-CL1-
C 2	198.4	99.2	450	1	AI720016	as55c01.x
C 3	198.4	99.2	456	1	AI818813	WK91F09.x
C 4	198.4	99.2	462	6	CA437924	UI-H-DHO-
C 5	198.4	99.2	471	1	AA972758	oq01f01.s
C 6	198.4	99.2	486	2	BF224310	7q75a08.x
C 7	198.4	99.2	509	6	CD673147	fg21g04.y
C 8	198.4	99.2	562	2	BF063923	7h90e01.x
C 9	198.4	99.2	601	5	BUS80378	in33b12.x
C 10	198.4	99.2	634	1	AV713706	AV713706
C 11	198.4	99.2	635	1	AV717017	AV717017
C 12	198.4	99.2	637	1	AV717339	AV717339
C 13	198.4	99.2	637	4	BM685551	UI-E-CIO-
C 14	198.4	99.2	640	5	BU681561	UI-CF-EC1
C 15	198.4	99.2	642	5	BM996167	UI-H-DHO-
C 16	198.4	99.2	655	5	BM994812	UI-H-DHO-
C 17	198.4	99.2	666	6	CD512640	AGENCOURT
C 18	198.4	99.2	699	5	BQ000814	UI-H-DHI-
C 19	198.4	99.2	709	6	CB957742	AGENCOURT
C 20	198.4	99.2	715	5	BU624736	UI-H-FGI-
C 21	198.4	99.2	720	1	AI971174	WR26a07.x
C 22	198.4	99.2	736	6	CD523234	AGENCOURT
C 23	198.4	99.2	753	6	CA449867	UI-H-E-CL1
C 24	196.8	98.4	357	2	BE858648	7f99g08.x

C 25	196.8	98.4	389	2	BE671641	BE671641 7a55a06.x
C 26	196.8	98.4	389	6	C06208	C06208 C06208 Huma
C 27	196.8	98.4	394	1	AI927247	AI927247 wn24c03.x
C 28	196.8	98.4	411	1	AA047236	AA047236 zk74a08.s
C 29	196.8	98.4	415	1	AA429323	AA429323 zw07b05.x
C 30	196.8	98.4	415	7	H99781	H99781 yx36g07.s1
C 31	196.8	98.4	419	5	BUS32858	BUS32858 AGENCOURT
C 32	196.8	98.4	432	1	AA804854	AA804854 of44g07.s
C 33	196.8	98.4	437	1	AA576253	AA576253 nm62b07.s
C 34	196.8	98.4	437	5	BU077671	BU077671 im36d12.y
C 35	196.8	98.4	439	1	AA024754	AA024754 zg76g09.s
C 36	196.8	98.4	444	1	AI927356	AI927356 wn49g02.x
C 37	196.8	98.4	449	2	BE466610	BE466610 h223a10.x
C 38	196.8	98.4	455	7	W52063	W52063 zc90a02.s1
C 39	196.8	98.4	461	1	AI086784	AI086784 oz77g03.x
C 40	196.8	98.4	462	1	AI298809	AI298809 qm91d11.x
C 41	196.8	98.4	463	1	AI075862	AI075862 oz25b08.x
C 42	196.8	98.4	469	1	AI298305	AI298305 qm92d03.x
C 43	196.8	98.4	479	5	BU680472	BU680472 UI-CF-EC1
C 44	196.8	98.4	480	4	BG149323	BG149323 nad26d10.
C 45	196.8	98.4	486	1	AI478647	AI478647 tm54f08.x

## ALIGNMENTS

RESULT 1  
BU729482/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BU729482 424 bp mRNA linear EST 09-OCT-2002  
UI-E-CL1-afe-p-03-0-UI.s1 UI-E-CL1 Homo sapiens cDNA clone  
UI-E-CL1-afe-p-03-0-UI 3', mRNA sequence.  
ACCESSION BU729482  
VERSION BU729482.1 GI:23652409  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 424)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-36, >AT richLow\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1..424  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-CL1-afe-p-03-0-UI"  
/tissue\_type="human retina"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-CL1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-Cl1 is a normalized cDNA library containing the following tissue(s): retina. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:8791-8806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7r3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the  $\phi$ T18 tail. The sequence tag for this library is CCGGG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG TISSUE=human retina TAG LIB=UI-E-Cl1 TAG\_SF0=CCGGG"

```

/db_xref="taxon:9608"
/clone="IMAGE:2332608"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/lab_clone_lib="Barstead colon HPU87B"
/notes="Organ: colon; Vector: pUT73D-Pac (Pharmacia) with a modified polynucleic acid (PNA) primer. The PNA primer strand cDNA was primed with a Not I - oligo(dT) primer. The cDNA was ligated to Eco RI adaptors (5' ATTCACTAGTGAAT 3', and 5' ATTCACTAGT 3'), digested with Not I and Eco RI sites of the modified pUT73 vector. Library constructed by Bob Barstead."
```

## ORIGIN

	Query Match	Best Local Similarity	99.2%;	Score 198.4;	DB 1;	Length 450;
	Matches 199;	Conservative	99.5%;	0;	No. 3.8e-31;	Indels 0; Gaps 0;
QY	1	TGCAGCTAACCTATATTTTCCCTAGCTTTCCCCAGACACCTTGCTTTTATTTATTAATG	60			
DB	345	TGCAGCTAACCTATATTTTCCCTAGCTTTCCCCAGACACCTTGCTTTTATTTATTAATG	286			
QY	61	AATTTTGTGTGATGTGAAACATTAATGCTTTAAGTAAATGTTCTTATTTAAGTTA	120			
DB	285	AATTTTGTGTGTGATGTGAAACATTAATGCTTTAAGTAAATGTTCTTATTTAAGTTA	226			
QY	121	TTGATGTTTTAAGTTTTATCTTTTCATGCTACTAGTGTTTTTTAGATACAGAGACTTTGGGA	180			
DB	225	TTGATGTTTTAAGTTTTATCTTTTCATGCTACTAGTGTTTTTTAGATACAGAGACTTTGGGA	166			
QY	181	AATTGCTTTTCTCTTTGTAC	200			
DB	165	AATTGCTTTTCTCTTTGAAC	146			

RESULT 3  
AI818813/c linear EST 07-MAR-2000  
LOCUS AI818813 456 bp mRNA  
W391F09.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2422793 3'  
DEFINITION W391F09.x1 NCI CGAP Lu19 Homo sapiens cDNA clone IMAGE:2422793 3'  
similar to gb:W24545 MONOCYTE CHEMOTACTIC PROTEIN 1 PRECURSOR  
(MCP-1). mRNA sequence.

(ROMANOV; MANDELSTAM)  
ACCESSION AI818813 GI:5437892  
VERSION AI818813.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Ceratobdini; Homnidae; Homo.

REFERENCE	1. (bases 1 to 456)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps@mail.nih.gov">cgaps@mail.nih.gov</a> Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <a href="http://www-bio.llnl.gov/bbrp/image/image.html">www-bio.llnl.gov/bbrp/image/image.html</a> Insert Length: 618 Std Error: 0.00 Seq primer: -40UP from Gibco Seq quality sequence stop: 438.

```

FEATURES
  source
    high quality seqs
    Location/Qualifiers
      1..456
        /organism="Homo sapiens"
        /mol type="mRNA"

```

db 183 AATGCTTTTCCTCTTGAAC 164

LOCUS AI720016 450 bp mRNA  
DEFINITION IMAGE5501.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone  
AI720016.1 precursor (human); mRNA sequence.  
ACCESSION AI720016  
VERSION AI720016.1 GI:5037272

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa;  
Mammalia; Eutheria;  
Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 450)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Tan, M., Martin, J., Moore, B., Schellenberg, K., Stenoe, M., Tan, F., Thesling, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST project

JOURNAL Unpublished (1997)

COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LIND; contact the  
LIND for further information.

IMAGE Consortium (info@image.lnli.gov) for further information.

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .450

FEATURES

source

```

/db_xref="taxon:9606"
/clone="IMAGE:2422793"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Lu19"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."

```

## ORIGIN

```

Query Match          99.2%; Score 198.4; DB 1; Length 456;
Best Local Similarity 99.5%; Pred. No. 3.8e-31;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TGCAGCTAACTATTTCCTAGCTTCCCGACACCTTGTTTATTATTATAATG 60
    |||||
Db 346 TGCAGCTAACTATTTCCTAGCTTCCCGACACCTTGTTTATTATTATAATG 287
    |||||

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTAATCTTATTAAATG 120
    |||||
Db 286 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTAATCTTATTAAATG 227
    |||||

QY 121 TTGATGTTTTAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
    |||||
Db 226 TTGATGTTTTAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 167
    |||||

QY 181 AATGCTTTTCCTCTGTAC 200
    |||||
Db 166 AATGCTTTTCCTCTGTAAAC 147
    |||||

```

## RESULT 4

```

CA437924/c
LOCUS
DEFINITION
CA437924
CA437924.1 GI:24802344
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 462)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-36, >AT richLow_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

```

## FEATURES

```

source
1. .462
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="UI-H-DHO-aur-1-19-0-UI"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DH0"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DH0 is a cDNA library containing the following
tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma
in Lung. The library was constructed according to Bonaldo,
Lennon and Soares, Genome Research, 6:791-806, 1996. First
strand cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is AGATCAATTGC.
TAG_TISSUE=Lung
TAG_LIB=UI-H-DHO
TAG_SEQ=AGATCAATTGC"

```

## ORIGIN

```

Query Match          99.2%; Score 198.4; DB 6; Length 462;
Best Local Similarity 99.5%; Pred. No. 3.8e-31;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTCCTAGCTTCCCGACACCTTGTTTATTATTATAATG 60
    |||||
Db 364 TGCAGCTAACTATTTCCTAGCTTCCCGACACCTTGTTTATTATTATAATG 305
    |||||

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTAATCTTATTAAATG 120
    |||||
Db 304 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTAATCTTATTAAATG 245
    |||||

QY 121 TTGATGTTTTAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
    |||||
Db 244 TTGATGTTTTAAGTTTATCTTTCATGCTAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 185
    |||||

QY 181 AATGCTTTTCCTCTGTAC 200
    |||||
Db 184 AATGCTTTTCCTCTGTAAAC 165
    |||||

```

## RESULT 5

```

AA972758/c
LOCUS
DEFINITION
AA972758
AA972758.1 GI:3147938
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 471)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

Tue Aug 9 17:01:25 2005

www.bio.lnl.gov/bbrp/image/image.html  
Insert Length: 541 Std Error: 0.00  
Seq primer: -40m13 fwd. RT from Amersham  
High quality sequence stop: 447.  
Location/Qualifiers

FEATURES  
source

1. .471  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1585081"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu5"  
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTT3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 99.2%; Score 198.4; DB 1; Length 471;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTATTTCCCTAGCTTCCCGACACCTGTTTATTATTATAATG 60  
Db 348 TGCAGCTAACTATTTCCCTAGCTTCCCGACACCTGTTTATTATTATAATG 289  
QY 61 AATTTGTTTGTGATGTAACATTATGCTTAAAGTAATGTTAACTTATTAAAGTTA 120  
Db 288 AATTTGTTTGTGATGTAACATTATGCTTAAAGTAATGTTAACTTATTAAAGTTA 229  
QY 121 TTGATGTTTAAAGTTAATCTTCTGTTAGTCTGTTTATTAGTACAGAGACTTGGGA 180  
Db 228 TTGATGTTTAAAGTTAATCTTCTGTTAGTCTGTTTATTAGTACAGAGACTTGGGA 169  
QY 181 AATGCTTTTCTCTCTGTAC 200  
Db 168 AATGCTTTTCTCTCTGTAC 149

RESULT 6  
BF224310/c

LOCUS  
DEFINITION  
BF224310 489 bp mRNA linear EST 30-MAR-2001  
7q75a08.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3703982 3,  
similar to SW:SY02\_HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2  
PRECUSOR 1, mRNA sequence.  
BF224310 GI:11131582  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 489)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.lnl.gov  
High quality sequence stop: 416.

FEATURES  
source

Location/Qualifiers  
1. .489  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3703982"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 99.2%; Score 198.4; DB 2; Length 489;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTATTTCCCTAGCTTCCCGACACCTGTTTATTATTATAATG 60  
Db 353 TGCAGCTAACTATTTCCCTAGCTTCCCGACACCTGTTTATTATTATAATG 294  
QY 61 AATTTGTTTGTGATGTAACATTATGCTTAAAGTAATGTTAACTTATTAAAGTTA 120  
Db 293 AATTTGTTTGTGATGTAACATTATGCTTAAAGTAATGTTAACTTATTAAAGTTA 234  
QY 121 TTGATGTTTAAAGTTAATCTTCTGTTAGTCTGTTTATTAGTACAGAGACTTGGGA 180  
Db 233 TTGATGTTTAAAGTTAATCTTCTGTTAGTCTGTTTATTAGTACAGAGACTTGGGA 174  
QY 181 AATGCTTTTCTCTCTGTAC 200  
Db 173 AATGCTTTTCTCTCTGTAC 154

RESULT 7  
CD673147

LOCUS  
DEFINITION  
CD673147 506 bp mRNA linear EST 24-JUN-2003  
fg21d04.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone  
CD673147 GI:32174878  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 506)  
Wistow, G., Bernstein, S.L., Ray, S., Wyatt, M.K., Behal, A.,  
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.  
Expressed sequence tag analysis of adult human iris for the NEIBank  
Project: steroid-response factors and similarities with retinal  
pigment epithelium  
Mol. Vis. 8 (4), 185-195 (2002)  
22103462  
12107412  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 21 row: d column: 04  
Seq primer: M13Rpl reverse primer (ABI).  
Location/Qualifiers  
1. .506

FEATURES  
source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="fg21d04"  
 /tissue\_type="Iris"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Iris cDNA (Normalized): fg"  
 /notes="Organ: Eye; Vector: pCMVSPORT6; A human iris library (bx) was normalized by self-subtraction. One portion of double stranded plasmid DNA representing the library was linearized by NotI. This NotI digested library was used as a template for biotinylated RNA synthesis using SP6 RNA polymerase. Another portion of the double stranded plasmid library was converted to single-stranded circles in vitro using Gene II and Exonuclease III (Life Technologies). Single-stranded DNA (1 mg) was hybridized (Cot 500) with 41 mg of Bio-RNA and vector blocking oligonucleotides. The hybridized Bio-RNA/ss-circles were removed by streptavidin:phenol extraction. EST analysis was performed on the library at the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 99.2%; Score 198.4; DB 6; Length 506;  
 Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
 Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTGTTTATTTATTAATG 60  
 |||||  
 Db 261 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTGTTTATTTATTAATG 320  
 |||||

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTTAAATCTTATTAAGTTA 120  
 |||||  
 Db 321 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTTAAATCTTATTAAGTTA 380  
 |||||

QY 121 TTGATGTTTTAAGTTATCTTTCATGCTAGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 180  
 |||||  
 Db 381 TTGATGTTTTAAGTTATCTTTCATGCTAGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 440  
 |||||

QY 181 AATTGCTTTTCCTCTGTGAC 200  
 |||||  
 Db 441 AATTGCTTTTCCTCTGTAAC 460  
 |||||

RESULT 8  
 BF063923/c  
 LOCUS  
 DEFINITION 7h90e01.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3323256 3' similar to SW:SY02 HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2  
 PRECURSOR ; mRNA sequence.

ACCESSION BF063923  
 VERSION BF063923.1 GI:10822833  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 562)  
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 TITLE Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL, send email to: info@image.llnl.gov

Seq primer: -40UP from Gibco  
 High quality sequence stop: 433.  
 Location/Qualifiers  
 1..562

## FEATURES

source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3323256"  
 /tissue\_type="colon tumor, RER+"  
 /lab\_host="PH10B"  
 /clone\_lib="NCI CGAP Col6"  
 /notes="Organ: colon; Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP Col6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1057416-1061255, and 1144584-1145351).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 99.2%; Score 198.4; DB 2; Length 562;  
 Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
 Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTGTTTATTTATTAATG 60  
 |||||  
 Db 346 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTGTTTATTTATTAATG 287  
 |||||

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTTAAATCTTATTAAGTTA 120  
 |||||  
 Db 286 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATGTTAAATCTTATTAAGTTA 227  
 |||||

QY 121 TTGATGTTTTAAGTTATCTTTCATGCTAGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 180  
 |||||  
 Db 226 TTGATGTTTTAAGTTATCTTTCATGCTAGTACTAGTGTGTTTTTAGATACAGAGACTTGGGGA 167  
 |||||

QY 181 AATTGCTTTTCCTCTGTGAC 200  
 |||||  
 Db 166 AATTGCTTTTCCTCTGTAAC 147  
 |||||

RESULT 9  
 BU580378/c  
 LOCUS  
 DEFINITION in33b12.x1 Human Fetal Pancreas 1B Homo sapiens cDNA 3' similar to SW:SY02 HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2 PRECURSOR ; mRNA sequence.

ACCESSION BU580378  
 VERSION BU580378.1 GI:23065638  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 601)  
 REFERENCE Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kastner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 COMMENT Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu

Tue Aug 9 17:01:25 2005

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@imgate.wustl.edu)  
Seq primer: -40UP from Gibco  
High quality sequence stop: 454.  
Location/Qualifiers

FEATURES  
source  
1. .601  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Fetal Pancreas (4 Pooled Donors, 18 - 20 weeks, Scratagene #738023)"  
/dev\_stage="Fetal Pancreas 1B"  
/clone\_lib="Human Fetal Pancreas 1B"  
/notes="Vector: pBluescript SK(-); Site 1: NotI; Site 2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size -1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN  
Query Match 99.2%; Score 198.4; DB 5; Length 601;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 60  
Db 348 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 289  
QY 61 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 120  
Db 288 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 229  
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 180  
Db 228 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 169  
QY 181 AATTGCTTTTCTCTTGTAC 200  
Db 168 AATTGCTTTTCTCTTGTAC 149

ORIGIN

Query Match 99.2%; Score 198.4; DB 5; Length 601;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 60  
Db 348 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 289  
QY 61 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 120  
Db 288 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 229  
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 180  
Db 228 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 169  
QY 181 AATTGCTTTTCTCTTGTAC 200  
Db 168 AATTGCTTTTCTCTTGTAC 149

RESULT 10

AV713706 634 bp mRNA linear EST 11-OCT-2000  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 634)  
Xing, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,  
Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,  
Cheng, Z., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Cheng, Z., and Han, Z.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

Location/Qualifiers  
1. .634  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DCBBYB12"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"  
/note="Vector: pTriplex2; Site\_1: sfIIA; Site\_2: sfIIB"

ORIGIN

Query Match 99.2%; Score 198.4; DB 1; Length 634;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 60  
Db 390 TGCAGCTAACTATTTTCCCTAGCTTCCCGACACCTTGTATTTATTTAATAAG 449  
QY 61 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 120  
Db 450 AATTTTGTGTTGATGCTGAACATTATGCTTAAGTAATGTTAATTTAATAAGTTA 509  
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 180  
Db 510 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTGTGTTTATAGATACAGACTTGGGGA 569  
QY 181 AATTGCTTTTCTCTTGTAC 200  
Db 570 AATTGCTTTTCTCTTGTAC 589

RESULT 11

AV717017 635 bp mRNA linear EST 16-OCT-2000  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 635)  
Xing, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,  
Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,  
Cheng, Z., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Cheng, Z., and Han, Z.

TITLE  
JOURNAL  
COMMENT  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES

source  
1. .635  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DCBBYB06"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"  
/note="Vector: pTriplex2; Site\_1: sfIIA; Site\_2: sfIIB"

ORIGIN

```
Query Match      99.2%; Score 198.4; DB 1; Length 635;
Best Local Similarity 99.5%; Pred. No. 3.8e-31;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTTGTTTATTTATTAATG 60
    |||
Db 395 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTTGTTTATTTATTAATG 454
    |||

QY 61 AATTTGTTGTTGATGTAACATTATGCTTAAAGTAATGTTAAATCTTATTTAAGTTA 120
    |||
Db 455 AATTTGTTGTTGATGTAACATTATGCTTAAAGTAATGTTAAATCTTATTTAAGTTA 514
    |||

QY 121 TTGATGTTTAACTTATCTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
    |||
Db 515 TTGATGTTTAACTTATCTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 574
    |||

QY 181 AATTGCTTTTCCTCTGTGAC 200
    |||
Db 575 AATTGCTTTTCCTCTTGAAC 594
    |||

RESULT 12
AV1717339 637 bp mRNA linear EST 16-OCT-2000
LOCUS
DEFINITION AV1717339 DCB Homo sapiens cDNA clone DCBXC09 5', mRNA sequence.
ACCESSION AV1717339
VERSION AV1717339.1 GI:10814491
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 637)
AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,S., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
Lu,G., Cheng,Z. and Han,Z.
TITLE Homo sapiens cDNA DCB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBXC09"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB"

ORIGIN
Query Match      99.2%; Score 198.4; DB 1; Length 637;
Best Local Similarity 99.5%; Pred. No. 3.8e-31;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTTGTTTATTTATTAATG 60
    |||
Db 395 TGCAGCTAACTTATTTCCCTAGCTTCCCGAGACACCTTGTTTATTTATTAATG 454
    |||

QY 61 AATTTGTTGTTGATGTAACATTATGCTTAAAGTAATGTTAAATCTTATTTAAGTTA 120
    |||
Db 455 AATTTGTTGTTGATGTAACATTATGCTTAAAGTAATGTTAAATCTTATTTAAGTTA 514
    |||

QY 121 TTGATGTTTAACTTATCTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 180
    |||
```

```
Db 515 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTGGGGA 574
    |||

QY 181 AATTGCTTTTCCTCTGTGAC 200
    |||
Db 575 AATTGCTTTTCCTCTTGAAC 594
    |||

RESULT 13
BM685551 637 bp mRNA linear EST 27-FEB-2002
LOCUS
DEFINITION UI-E-C10-aae-g-12-0-UI.r1 UI-E-C10 Homo sapiens cDNA clone
UI-E-C10-aae-g-12-0-UI 5', mRNA sequence.
ACCESSION BM685551
VERSION BM685551.1 GI:18995447
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 637)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 491-531, >AT-rich#Low_complexity (matched complement)
Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C10-aae-g-12-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C10"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C10 is a cDNA library containing the following
tissue(s): RPE and Choroid. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is ACCTA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

ORIGIN
Query Match      99.2%; Score 198.4; DB 4; Length 637;
Best Local Similarity 99.5%; Pred. No. 3.8e-31;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

us-09-899-276c-2.rst

Tue Aug 9 17:01:25 2005

Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is

AGTGGCTTAC.

TAG\_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371 and 380-383

TAG\_LIB=UI-CF-EC1

TAG\_SEQ=AAAGTGCTTAC"

ORIGIN

Query Match 99.2%; Score 198.4; DB 5; Length 640;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTTATTAATG 60  
Db 364 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTTATTAATG 305

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATCTTAATCTTTATTAAGTTA 120  
Db 304 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATCTTAATCTTTATTAAGTTA 245

QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTCTTTTATGATACAGAGACTTGGGA 180  
Db 244 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTCTTTTATGATACAGAGACTTGGGA 185

QY 181 AATGCTTTTCCCTCTTGAC 200  
Db 184 AATGCTTTTCCCTCTTGAC 165

RESULT 15

BM996167 642 bp mRNA linear EST 17-JUN-2002  
LOCUS UI-H-DH0-aun-d-15-0-UI.s1 NCI CGAP\_DHO Homo sapiens cDNA clone  
DEFINITION IMAGE:5871710 3', mRNA sequence.

ACCESSION BM996167 GI:19721068  
VERSION BM996167  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 642)  
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
The following repetitive elements were found in this cDNA sequence: 1-37, >AT rich#Low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes

FEATURES

Location/Qualifiers

1..642  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5871710"  
/tissue\_type="Metastatic Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI\_CGAP\_DHO"

QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTTATTAATG 60  
Db 164 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACCTTGTATTTTATTTATTAATG 223

QY 61 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATCTTAATCTTTATTAAGTTA 120  
Db 224 AATTTGTTTGTGATGGAACATTATGCTTAAGTAATCTTAATCTTTATTAAGTTA 283

QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTCTTTTATGATACAGAGACTTGGGA 180  
Db 284 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTAGTCTTTTATGATACAGAGACTTGGGA 343

QY 181 AATGCTTTTCCCTCTTGAC 200  
Db 344 AATGCTTTTCCCTCTTGAC 363

RESULT 14

BU681561 640 bp mRNA linear EST 07-OCT-2002  
LOCUS UI-CF-EC1-acj-k-08-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone  
DEFINITION UI-CF-EC1-acj-k-08-0-UI 3', mRNA sequence.

ACCESSION BU681561 GI:23531598  
VERSION BU681561  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 640)  
REFERENCE Bonaldo,M.P., Lennon,G. and Soares,M.B.  
AUTHORS Normalization and subtraction: two approaches to facilitate gene  
TITLE discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: McCray, PB  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: [paul-mccray@uiowa.edu](mailto:paul-mccray@uiowa.edu)  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics ([www.resgen.com](http://www.resgen.com)) or from Open Biosystems ([www.openbiosystems.com](http://www.openbiosystems.com)).  
The following repetitive elements were found in this cDNA sequence: 1-37, >AT rich#Low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes

FEATURES

Location/Qualifiers

1..640  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-EC1-acj-k-08-0-UI"  
/tissue\_type="Lung"  
/dev\_stage="Adult and Fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-EC1"  
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a

/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DH0 is a cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATTGC.  
TAG TISSUE=lung  
TAG\_LIB=UI-H-DH0  
TAG\_SEQ=AGATCATTGC"

ORIGIN

Query Match 99.2%; Score 198.4; DB 5; Length 642;  
Best Local Similarity 99.5%; Pred. No. 3.8e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 TGCAGCTAACTATTTCCTTAGCTTCCCGACACACCTTGTATTATTATAATG 60  
|||  
Db 364 TGCAGCTAACTATTTCCTTAGCTTCCCGACACACCTTGTATTATTATAATG 305  
|||  
  
QY 61 AATTTGTTTGTGATGTGAACATTATGCCTTAAGTAATGTTAATCTTATTAAAGTTA 120  
|||  
Db 304 AATTTGTTTGTGATGTGAACATTATGCCTTAAGTAATGTTAATCTTATTAAAGTTA 245  
|||  
  
QY 121 TTGATGTTTTAAGTTTATCTTTCATGCTAGTCTGTTTTTTAGATACAGAGACTTCGGGA 180  
|||  
Db 244 TTGATGTTTTAAGTTTATCTTTCATGCTAGTCTGTTTTTTAGATACAGAGACTTCGGGA 185  
|||  
  
QY 181 AATTGCTTTTCCTCTTGATC 200  
|||  
Db 184 AATTGCTTTTCCTCTTGATC 165  
|||

Search completed: August 4, 2005, 14:32:05  
Job time : 951.968 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 41.2146 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-2

Perfect score: 200

Sequence: 1 tgcagtaactattttccc.....aattgtttctctgtac 200

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCrUS COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	198.4	99.2	772	4	US-08-437-306-1
2	196.8	98.4	647	4	US-09-777-430C-72
3	196.8	98.4	725	4	US-09-591-992-3
4	196.8	98.4	725	4	US-09-023-655-1344
5	196.8	98.4	741	3	US-07-927-391-25
6	196.8	98.4	741	4	US-09-023-655-1165
7	196.8	98.4	752	6	5212073-1
8	196.8	98.4	752	6	5212073-1
9	196.8	98.4	1712	4	US-09-148-545-106
10	196.8	98.4	1822	4	US-09-148-545-105
11	195.2	97.6	661	4	US-09-814-915A-104
12	45.2	22.6	840	4	US-09-016-434-1033
13	43.6	21.8	612	4	US-09-502-540-1357
14	43.2	21.6	585	4	US-09-328-352-761
15	43.2	21.6	346112	4	US-09-949-016-13165
16	42.2	21.1	6124	3	US-08-213-419B-3
17	41	20.5	1738	4	US-09-918-909A-27
18	40.8	20.4	285478	4	US-09-949-016-13362
19	40	20.0	396	4	US-09-840-173-10
20	40	20.0	396	4	US-09-713-550-10
21	40	20.0	396	4	US-09-825-294-10
22	40	20.0	396	4	US-09-970-966-10
23	39.4	19.7	6409	4	US-09-967-908A-1
24	39.4	19.7	6409	4	US-10-159-151-1
25	39.2	19.6	49440	4	US-09-949-016-14150
26	39	19.5	601	4	US-09-949-016-182221
27	39	19.5	133358	4	US-09-949-016-16964

Sequence 12651, A  
Sequence 1280, Ap  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 13, Appli  
Sequence 135597,  
Sequence 15546, A  
Sequence 17037, A  
Sequence 12106, A  
Sequence 1, Appli  
Sequence 10, Appli  
Sequence 10, Appli  
Sequence 13248, A  
Sequence 98, Appli  
Sequence 98, Appli  
Sequence 98, Appli

28 39 19.5 133360 4 US-09-949-016-12651  
29 38.8 19.4 1039 4 US-09-902-540-1280  
30 38.8 19.4 2246 3 US-09-363-708-3  
31 38.8 19.4 2246 4 US-09-083-587-3  
32 38.8 19.4 19124 2 US-08-487-828B-13  
33 38.4 19.2 601 4 US-09-949-016-135597  
34 38.4 19.2 187595 4 US-09-949-016-15546  
35 38.4 19.2 260286 4 US-09-949-016-17037  
36 38.4 19.2 260293 4 US-09-949-016-12106  
37 38.2 19.1 2674 3 US-09-817-180-1  
38 38.2 19.1 2674 4 US-10-003-295-1  
39 38 19.0 1051 3 US-09-245-041-10  
40 38 19.0 1051 4 US-09-358-055B-10  
41 38 19.0 1051 4 US-09-893-238-10  
42 38 19.0 12313 4 US-09-949-016-13248  
43 37.8 18.9 208 2 US-08-967-101-98  
44 37.8 18.9 208 2 US-08-592-541-98  
45 37.8 18.9 208 3 US-09-124-698-98

#### ALIGNMENTS

RESULT 1  
US-08-437-306-1  
; Sequence 1, Application US/08437306  
; Patent No. 6787645  
; GENERAL INFORMATION:  
; APPLICANT: Rollins, Barrett J.  
; APPLICANT: Stiles, Charles D.  
; APPLICANT: Wong, Gordon G.  
; TITLE OF INVENTION: No. 6787645el Human Cytokine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/437,306  
; FILING DATE: 09-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/228,931  
; FILING DATE: 13-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/003,136  
; FILING DATE: 12-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,515  
; FILING DATE: 16-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/351,008  
; FILING DATE: 12-MAY-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: DPCI-196A2Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 772 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..369
; US-08-437-306-1

Query Match 99.2%; Score 198.4; DB 4; Length 772;
Best Local Similarity 99.5%; Pred. No. 9.6e-39;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTTTATTAATG 60
DB 398 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTTTATTAATG 457
QY 61 AATTTTGTGTGTGATGTGAACAATATGCTTAAAGTAATGTTAATCTTATTTAAGTTA 120
DB 458 AATTTTGTGTGTGATGTGAACAATATGCTTAAAGTAATGTTAATCTTATTTAAGTTA 517
QY 121 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTATTTTATAGATACAGAGACTTGGGA 180
DB 518 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTATTTTATAGATACAGAGACTTGGGA 577
QY 181 AATTGCTTTTCCTCTGTGATC 200
DB 578 AATTGCTTTTCCTCTGTAAC 597

RESULT 2
US-09-777-430C-72
; Sequence 72, Application US/09777430C
; Patent No. 6780982
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Skrzypczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Teeska
; APPLICANT: Neir, Bruce P.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: FORS-04912
; CURRENT APPLICATION NUMBER: US/09/777,430C
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 647
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; US-09-777-430C-72

Query Match 98.4%; Score 196.8; DB 4; Length 647;
Best Local Similarity 94.0%; Pred. No. 2.3e-38;
Matches 108; Conservative 90; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTTTATTAATG 60
DB 326 UGCAGCUAAUUUUUCCUCCUAGCUUUUCCCGACACACCCUGUUUUUUUUUUUAUUAUG 385
QY 61 AATTTTGTGTGTGATGTGAACAATATGCTTAAAGTAATGTTAATCTTATTTAAGTTA 120
DB 386 AAUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUA 445
QY 121 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTATTTTATAGATACAGAGACTTGGGA 180
DB 446 UUGAUGUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 505
QY 181 AATTGCTTTTCCTCTGTGATC 200
DB 506 AATGCTTTTCCTCTGTAAC 525

```

```
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1344:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 725 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: G34513
/ US-09-023-655-1344

Query Match          98.4%; Score 196.8; DB 4; Length 725;
Best Local Similarity 99.0%; Pred. No. 2.3e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACACCTGTTTATTATTATAATG 60
Db      379 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACACCTGTTTATTATTATAATG 438

QY      61 AATTTGTTGTTGATGTGAACATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 120
Db      439 AATTTGTTGTTGATGTGAACATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 498

QY      121 TTGATGTTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTTTTAGATACAGAGACTTGGGGA 180
Db      499 TTGATGTTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTTTTAGATACAGAGACTTGGGGA 558

QY      181 AATTGCTTTTCTCTTGATC 200
Db      559 AATTGCTTTTCTCTTGATC 578

RESULT 5
US-09-023-391-25
/ Sequence 25, Application US/07927391
/ Patent No. 6001649
/ GENERAL INFORMATION:
/ APPLICANT: CAPUT, Daniel
/ APPLICANT: FERRARA, Pascual
/ APPLICANT: MILOUK, Brigitte
/ APPLICANT: MINTY, Adrian
/ APPLICANT: VITA, Natalio
/ TITLE OF INVENTION: Protein having a cytokin type
/ TITLE OF INVENTION: activity, and recombinant DNA, expression vector and hosts
/ TITLE OF INVENTION: for its preparation.
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: POLY & LARDNER
/ STREET: King Street Station, Suite 500, 1800 Diagonal
/ STREET: Road, PO Box 299
/ CITY: ALEXANDRIA
/ STATE: VIRGINIA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/07/927,391
/ FILING DATE: 19920929
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Saxe, Bernhard D
/ REGISTRATION NUMBER: 28,665
/ REFERENCE/DOCKET NUMBER: 16781/369
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-9300
/ TELEFAX: (703) 683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 25:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 741 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-07-927-391-25

Query Match          98.4%; Score 196.8; DB 3; Length 741;
Best Local Similarity 99.0%; Pred. No. 2.3e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACACCTGTTTATTATTATAATG 60
Db      395 TGCAGCTAACTTATTTTCCCTAGCTTTCCCGACACACCTGTTTATTATTATAATG 454

QY      61 AATTTGTTGTTGATGTGAACATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 120
Db      455 AATTTGTTGTTGATGTGAACATTATGCTTAAAGTAATGTTAAATCTTATTAAAGTTA 514

QY      121 TTGATGTTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTTTTAGATACAGAGACTTGGGGA 180
Db      515 TTGATGTTTTAAAGTTATCTTTCATGCTACTAGTGTGTTTTTTAGATACAGAGACTTGGGGA 574

QY      181 AATTGCTTTTCTCTTGATC 200
Db      575 AATTGCTTTTCTCTTGATC 594

RESULT 6
US-09-023-655-1165
/ Sequence 1165, Application US/09023655
/ Patent No. 6607879
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ TITLE OF INVENTION: EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN JE CYTOKINE
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1165:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 741 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g187434
; US-09-023-655-1165

Query Match 98.4%; Score 196.8; DB 4; Length 741;
Best Local Similarity 99.0%; Pred. No. 2.3e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 60
Db TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 454
QY 61 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 120
Db 455 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 514
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 180
Db 515 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 574
QY 181 AATGCTTTTCCCTCTGTAC 200
Db 575 AATGCTTTTCCCTCTTGAAC 594

RESULT 7
5212073-1
; Patent No. 5212073
; APPLICANT: ROLLINS, BARRETT; STILES, CHARLES; WONG, GORDON G.
; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN JE CYTOKINE
; NUMBER OF SEQUENCES: 1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/351,008
; FILING DATE: 12-MAY-1989
; SEQ ID NO: 1
; LENGTH: 752
5212073-1
Query Match 98.4%; Score 196.8; DB 6; Length 752;
Best Local Similarity 99.0%; Pred. No. 2.3e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 60
Db 388 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 447
QY 61 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 120
Db 448 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 507
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 180
Db 508 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 567
QY 181 AATGCTTTTCCCTCTGTAC 200
Db 568 AATGCTTTTCCCTCTTGAAC 587

RESULT 8
5212073-1
; Patent No. 5212073
; APPLICANT: ROLLINS, BARRETT; STILES, CHARLES; WONG, GORDON G.
; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN JE CYTOKINE
; NUMBER OF SEQUENCES: 1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/351,008
; FILING DATE: 12-MAY-1989
; SEQ ID NO: 1
; LENGTH: 752
5212073-1
Query Match 98.4%; Score 196.8; DB 6; Length 752;
Best Local Similarity 99.0%; Pred. No. 2.3e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 60
Db 388 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTATTTTATTTAATG 447
QY 61 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 120
Db 448 AATTTGTTTGTGATGTAACATTAATGCTTAAAGTAATGTTAAATCTTAAAGTTA 507
QY 121 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 180
Db 508 TTGATGTTTAAAGTTATCTTTTCATGCTAGTACTGTTTATAGATACAGAGCTTGGGA 567
QY 181 AATGCTTTTCCCTCTGTAC 200
Db 568 AATGCTTTTCCCTCTTGAAC 587

RESULT 9
US-09-148-545-106
; Sequence 106, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001PI
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US99/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
```



us-09-899-276c-2.rn1

Tue Aug 9 17:01:25 2005

```

; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 1712

Query Match      98.4%; Score 196.8; DB 4; Length 1712;
Best Local Similarity 99.0%; Pred. No. 2.6e-38;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTTATTTTATTTATG 60
Db      1315 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTTATTTTATTTATG 1374

QY      61  AATTTGTTTCTCATGTGAACATTATGCTTAAAGTAAATGTTAAATCTTATTTAGTTA 120
Db      1375 AATTTGTTTCTCATGTGAACATTATGCTTAAAGTAAATGTTAAATCTTATTTAGTTA 1434

QY      121 TTGATGTTTAAAGTTTATCTTTCAAGTAACTAGTGTGTTTTAGATACAGAGACTTCGGGA 180
Db      1435 TTGATGTTTAAAGTTTATCTTTCAAGTAACTAGTGTGTTTTAGATACAGAGACTTCGGGA 1494

QY      181 AATGCTTTCTCTCTGTAC 200
Db      1495 AATGCTTTCTCTCTGTAC 1514

RESULT 10
US-09-148-545-105
; Sequence 105, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,502
```

EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 05-Sep-1997  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881

EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 280  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 105  
LENGTH: 1822

Query Match 98.4%; Score 196.8; DB 4; Length 1822;  
Best Local Similarity 99.0%; Pred. No. 2.6e-38;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTATAATG 60  
1425 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTATAATG 1484

QY 61 AATTTTGTGTGATGTGAACATTAATGCTTAAGTAATGTAATTTTATTATAATG 120  
1485 AATTTTGTGTGATGTGAACATTAATGCTTAAGTAATGTAATTTTATTATAATG 1544

QY 121 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTATTTAGATACAGAGACTTCGGGA 180  
1545 TTGATGTTTAAAGTTATCTTTCATGTAAGTACTAGTGTATTTAGATACAGAGACTTCGGGA 1604

QY 181 AATTGCTTTCCCTGTGAC 200  
1605 AATTGCTTTCCCTGTGAC 1624

RESULT 11  
US-09-814-915A-104  
Sequence 104, Application US/09814915A  
Patent No. 6750015  
GENERAL INFORMATION:  
APPLICANT: Horwitz, Kathryn  
APPLICANT: Richer, Jennifer  
TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Rel.  
TITLE OF INVENTION: Theteto  
FILE REFERENCE: 2848-39  
CURRENT APPLICATION NUMBER: US/09/814,915A  
CURRENT FILING DATE: 2002-03-21  
PRIOR APPLICATION NUMBER: 60/214,870  
PRIOR FILING DATE: 2000-06-28  
NUMBER OF SEQ ID NOS: 108  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 104  
LENGTH: 661  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-814-915A-104

Query Match 97.6%; Score 195.2; DB 4; Length 661;  
Best Local Similarity 98.5%; Pred. No. 5.5e-38;  
Matches 197; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTATAATG 60  
357 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACACCTTGTATTTTATTATAATG 416

us-09-899-276c-2.rni

Tue Aug 9 17:01:25 2005

QY 61 AATTTTGTGGTGAATGAAACATATGCTTAAAGTAAATCTTATTAAGTTA 120  
 Db 417 AATTTTGTGGTGAATGAAACATATGCTTAAAGTAAATCTTATTAAGTTA 476  
 QY 121 TTGATGTTTAAATTAATCTTCAATGCTAGTGTGTTTAAAGTAAATCTTATTAAGTTA 180  
 Db 477 TTGATGTTTAAATTAATCTTCAATGCTAGTGTGTTTAAAGTAAATCTTATTAAGTTA 536  
 QY 181 AATGCTTTTCTCTTGTAC 200  
 Db 537 AATGCTTTTCTCTTGTAC 556

RESULT 12

US-09-016-434-1033

Sequence 1033, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION

NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016.434

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1033:

SEQUENCE CHARACTERISTICS:

LENGTH: 840 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BRSTNOT05

CLONE: 965517

US-09-016-434-1033

Query Match 22.6%; Score 45.2; DB 4; Length 840;

Best Local Similarity 66.9%; Pred. No. 0.049;

Matches 79; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 82 ACATATGCTTAAAGTAAATCTTAAATTAAGTAAATCTTAAAGTAAATCTT 141

Db 477 ATAATAATCTTAAAGTAAATCTTAAATTAAGTAAATCTTAAAGTAAATCT 536

QY 142 TCATGG-TACTAGTGTGTTTAAAGTAAATCTTAAAGTAAATCTTAAAGTAAATCT 198

Db 537 TCATACATCTTAAAGTAAATCTTAAAGTAAATCTTAAAGTAAATCTTAAAGTAAATCT 594

RESULT 13

US-09-902-540-1357/c

Sequence 1357, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 1357

LENGTH: 612

TYPE: DNA

ORGANISM: Myxococcus xanthus

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(612)

OTHER INFORMATION: unsure at all n locations

US-09-902-540-1357

Query Match 21.8%; Score 43.6; DB 4; Length 612;

Best Local Similarity 54.8%; Pred. No. 0.11;

Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 7 TAACTTATTTTCCCTAGCTTTCCCGACACCTTGTGTTTATTTTATTAATAAGTAAATTTT 66

Db 315 TANTTTTTTTTTTTTAAATTTTTTTTTTAAATATTTATTTTATTTTATTTTATTTTATTTT 256

QY 67 GTTGTGATGCTGAACATATGCTTAAAGTAAATGTTAAATCTTATTAAGTAAATGATG 126

Db 255 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 196

QY 127 TTTTAAAGTTTATCTTTCATGAGTACTAGTGTGTTT 161

Db 195 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 161

RESULT 14

US-09-328-352-761

Sequence 761, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 761

LENGTH: 585

TYPE: DNA

ORGANISM: Acinetobacter baumannii

US-09-328-352-761

Query Match 21.6%; Score 43.2; DB 4; Length 585;

Best Local Similarity 56.2%; Pred. No. 0.14;

Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 40 TTGTTTTTATTTTATTAATAAGTAAATTTTGTGTTGATGTGAACATTTATGCTTAAAGTAA 99

Db 320 TTATTTATTTTGGTATCATCATTTATGCTATTTAATAGGACCGTTTTTGGTGGTATTT 379

QY 100 TGTAAATCTTATTTAAAGTTATTAAGTTTATCTTTCAATGCTACTAGTGTGTTT 159

Db 380 TGTATTTTATTTAAAGATTTCTAATGTTAAATTTCAATTTTTTTTAGCAATTTCTATTTT 439

QY 160 TTAGATACAGAGACTTGGGAAAT 183  
||| ||| ||| |||  
Db 440 TTACTTAGGATGTATAGTTCAT 463

RESULT 15  
US-09-949-016-13165  
; Sequence 13165, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13165  
; LENGTH: 346112  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(346112)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13165

Query Match 21.8%; Score 43.2; DB 4; Length 346112;  
Best Local Similarity 56.2%; Pred. No. 0.36;  
Matches 81; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 28 TCCCCAGACACCTTGTATTTATTTATTAATGAATTTGTTGTGATGTGAACATTA 87  
Db 167788 TACCACACTACTTTCTTAAATTAATATACTCAATCTGACAGTTTGGAGAAAGAGTA 167847  
QY 88 TGCCTTAAGTAAGTTAAATCTTTATTTAAGTTATGATGTTTAAAGTTATCTTTCATGG 147  
Db 167848 AGAGTTACTACATGGAATGCCCTATTTAAGTTATATTGAGTTAGTTTGGCTGCATTT 167907  
QY 148 TACTAGTGTTTTATAGATACAGAG 171  
Db 167908 TTGCCTGGGTTTAAAGTATTGAG 167931

Search completed: August 4, 2005, 14:40:35  
Job time : 43.2146 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 356.761 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-2

Perfect score: 200

Sequence: 1 tgcagtaactattttccc.....aattgtttctctgtgac 200

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200	100.0	200	9	US-09-899-276-2
2	198.4	99.2	489	10	US-09-918-995-6347
3	198.4	99.2	1005	16	US-10-133-013-211
4	198.4	99.2	1472	16	US-10-247-671-41
5	198.4	99.2	11793	19	US-10-685-705-4
6	196.8	98.4	360	9	US-09-864-761-5454
7	196.8	98.4	508	19	US-10-775-169-294

8	196.8	98.4	647	9	US-09-777-430A-72	Sequence 72, Appl
9	196.8	98.4	647	22	US-10-875-094-72	Sequence 72, Appl
10	196.8	98.4	725	16	US-10-210-120-66	Sequence 66, Appl
11	196.8	98.4	725	18	US-10-641-643-1344	Sequence 1344, Ap
12	196.8	98.4	725	19	US-10-283-975A-58	Sequence 58, Appl
13	196.8	98.4	725	19	US-10-764-649-17	Sequence 17, Appl
14	196.8	98.4	725	19	US-10-445-790-3	Sequence 3, Appl
15	196.8	98.4	725	21	US-10-852-335A-46	Sequence 46, Appl
16	196.8	98.4	725	21	US-10-852-335A-54	Sequence 54, Appl
17	196.8	98.4	725	22	US-10-909-035-66	Sequence 66, Appl
18	196.8	98.4	739	17	US-10-170-385-396	Sequence 396, App
19	196.8	98.4	739	17	US-10-172-118-849	Sequence 849, App
20	196.8	98.4	739	18	US-10-342-887-849	Sequence 849, App
21	196.8	98.4	739	20	US-10-473-974-96	Sequence 96, Appl
22	196.8	98.4	739	21	US-10-278-698-35	Sequence 35, Appl
23	196.8	98.4	739	21	US-10-278-698-549	Sequence 549, App
24	196.8	98.4	741	18	US-10-641-643-1165	Sequence 1165, Ap
25	196.8	98.4	743	19	US-10-685-705-2	Sequence 2, Appl
26	196.8	98.4	756	16	US-10-133-013-210	Sequence 210, App
27	196.8	98.4	757	15	US-10-339-778-1	Sequence 1, Appl
28	196.8	98.4	757	17	US-10-388-360-377	Sequence 377, App
29	196.8	98.4	757	19	US-10-685-705-1	Sequence 1, Appl
30	196.8	98.4	757	20	US-10-733-878-516	Sequence 516, App
31	196.8	98.4	783	16	US-10-133-013-209	Sequence 209, App
32	196.8	98.4	804	10	US-09-971-429B-18	Sequence 18, Appl
33	196.8	98.4	804	13	US-10-044-090-583	Sequence 583, App
34	196.8	98.4	1005	16	US-10-133-013-211	Sequence 211, App
35	196.8	98.4	1102	14	US-10-198-846-9822	Sequence 9822, Ap
36	196.8	98.4	1472	16	US-10-247-671-41	Sequence 41, Appl
37	196.8	98.4	1712	9	US-09-981-876-106	Sequence 106, App
38	196.8	98.4	1712	10	US-09-148-545-106	Sequence 106, App
39	196.8	98.4	1822	9	US-09-981-876-105	Sequence 105, App
40	196.8	98.4	1822	10	US-09-148-545-105	Sequence 105, App
41	196.8	98.4	2776	22	US-10-833-656-2	Sequence 2, Appl
42	196.8	97.7	483	9	US-09-796-692-8779	Sequence 8779, Ap
43	195.4	97.7	483	14	US-10-040-862-8779	Sequence 8779, Ap
44	195.4	97.7	483	17	US-10-057-475B-8779	Sequence 8779, Ap
45	195.4	97.7	483	17	US-10-154-884B-8779	Sequence 8779, Ap

ALIGNMENTS

RESULT 1  
US-09-899-276-2  
; Sequence 2, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Dellus, Hajo  
; APPLICANT: Pouatka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 01627-023  
; CURRENT APPLICATION NUMBER: US/09/899,276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 200  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-2

Query Match 100.0%; Score 200; DB 9; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2e-32;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-09-899-276c-2.rnpb

Tue Aug 9 17:01:25 2005

```

QY 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 60
Db 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 60
QY 61 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 120
Db 61 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 120
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 180
Db 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 180
QY 181 AATTGCTTTTCCCTCTGTAC 200
Db 181 AATTGCTTTTCCCTCTGTAC 200

```

```

RESULT 2
US-09-918-995-6347
; Sequence 6347, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6347
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(489)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-6347

```

```

Query Match 99.2%; Score 198.4; DB 10; Length 489;
Best Local Similarity 99.5%; Pred. No. 5.4e-32; Indels 0; Gaps 0;
Matches 199; Conservative 0; Mismatches 1;
QY 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 60
Db 139 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 198
QY 61 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 120
Db 199 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 258
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 180
Db 259 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 318
QY 181 AATTGCTTTTCCCTCTGTAC 200
Db 319 AATTGCTTTTCCCTCTGTAC 338

```

```

RESULT 3
US-10-133-013-211/c
; Sequence 211, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astronomoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE

```

```

; FILE REFERENCE: PA-0049 US
; CURRENT APPLICATION NUMBER: US/10/133,013
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO 211
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 989010.5
US-10-133-013-211

```

```

Query Match 99.2%; Score 198.4; DB 16; Length 1005;
Best Local Similarity 99.5%; Pred. No. 6.6e-32; Indels 0; Gaps 0;
Matches 199; Conservative 0; Mismatches 1;
QY 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 60
Db 347 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 288
QY 61 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 120
Db 287 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 228
QY 121 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 180
Db 227 TTGATGTTTAAAGTTTATCTTTCATGCTACTAGTCTTTTATAGATACAGAGACTTGGGA 168
QY 181 AATTGCTTTTCCCTCTGTAC 200
Db 167 AATTGCTTTTCCCTCTGTAC 148

```

```

RESULT 4
US-10-247-671-41/c
; Sequence 41, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 1472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 989010.6
US-10-247-671-41

```

```

Query Match 99.2%; Score 198.4; DB 16; Length 1472;
Best Local Similarity 99.5%; Pred. No. 7.3e-32; Indels 0; Gaps 0;
Matches 199; Conservative 0; Mismatches 1;
QY 1 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 60
Db 347 TGCAGCTAACTTATTTTCCCTAGCTTCCCGACACCTTGTTTATTTTATTAATG 288
QY 61 AATTTTGTGTGTGATGTGAACATTATGCGCTTAAGTAAGTAACTTATTTAAGTTA 120

```

Db 287 AATTGTTTGTGATGTAACATTAATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 228  
121 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 180  
227 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 168  
QY 181 AATTGCTTTTCCCTCTGTAC 200  
Db 167 AATTGCTTTTCCCTCTGTAAC 148  
RESULT 5  
US-10-685-705-4  
; Sequence 4, Application US/10685705  
; Publication No. US20040177387A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: JAYAKRISHNA, Ambati  
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
; TITLE OF INVENTION: Degeneration  
; FILE REFERENCE: 050229-0415  
; CURRENT APPLICATION NUMBER: US/10/685,705  
; CURRENT FILING DATE: 2003-10-16  
; PRIOR APPLICATION NUMBER: 60/422,096  
; PRIOR FILING DATE: 2002-10-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 11793  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-685-705-4  
Query Match 99.2%; Score 198.4; DB 19; Length 11793;  
Best Local Similarity 99.5%; Pred. No. 1.3e-31;  
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCCAGACACCTTGTTTATTATTATTATAAG 60  
Db 9099 TGCAGCTAACTTATTTTCCCTAGCTTTCCCCAGACACCTTGTTTATTATTATAAG 9158  
QY 61 AATTGTTGTTGATGTAACATTAATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 120  
Db 9159 AATTGTTGTTGATGTAACATTAATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 9218  
QY 121 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 180  
Db 9219 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 9278  
QY 181 AATTGCTTTTCCCTCTGTAC 200  
Db 9279 AATTGCTTTTCCCTCTGTAAC 9298  
RESULT 6  
US-09-864-761-5454/c  
; Sequence 5454, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 5454  
; LENGTH: 360  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC005549.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5  
US-09-864-761-5454  
Query Match 98.4%; Score 196.8; DB 9; Length 360;  
Best Local Similarity 99.0%; Pred. No. 1.1e-31;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTTATTTTCCCTAGCTTTCCCCAGACACCTTGTTTATTATTATAAG 60  
Db 227 TGCAGCTAACTTATTTTCCCTAGCTTTCCCCAGACACCTTGTTTATTATTATAAG 168  
QY 61 AATTGTTGTTGATGTAACATTAATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 120  
Db 167 AATTGTTGTTGATGTAACATTAATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 108  
QY 121 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 180  
Db 107 TTGATGTTTAAAGTTTATCTTTCATGTAAGTCTGTTTATAGATACAGAGACTTGGGA 48  
QY 181 AATTGCTTTTCCCTCTGTAC 200  
Db 47 AATTGCTTTTCCCTCTGTAAC 28  
RESULT 7  
US-10-775-169-294  
; Sequence 294, Application US/10775169

```

; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; APPLICANT: Monitoring Drug Activities In Vivo
; TITLE OF INVENTION: Method for
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 294
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33)..(199)
; OTHER INFORMATION: Each "n" represents a nucleotide selected from a, t, g or c, or
; OTHER INFORMATION: contains no nucleotide.
; US-10-775-169-294

```

	Query Match	Best Local Similarity	Score 196.8;	DB 19;	Length 508;
	Matches 196;	Conservative	0;	Mismatches	2; Indels 0; Gaps 0;
			Pred. No. 1.2e-31;		
Qy	1	TCGAGCTAACTATTATTTCCCTAGCTTCCCCAGACACCTGTTTCTTTTATTATTATAATG	60		
Db	256	TCGAGCTAACTATTATTTCCCTAGCTTCCCCAGACACCTGTTTCTTTTATTATTATAATG	315		
Qy	61	AATTTGTTTGTGATGTGAACATTATGCCCTTAAGTAATGTTAATTCATTATTAAAGTTA	120		
Db	316	AATTTGTTTGTGATGTGAACATTATGCCCTTAAGTAATGTTAATTCATTATTAAAGTTA	375		
Qy	121	TTGATGTTTTAAAGTTATCTTTTCATGGTACTAGTGTTTTATAGATACACAGACTTTGGGGA	180		
Db	376	TTGATGTTTTAAAGTTATCTTTTCATGGTACTAGTGTTTTATAGATACACAGACTTTGGGGA	435		
Qy	181	AATTGCTTTTCCCTCTTGATC	200		
Db	436	AATTGCTTTTCCCTCTTGAAAC	455		

```

RESULT 8
US-09-777-430A-72
; Sequence 72, Application US/09777430A
; Patent No. US20020128465A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Skrzypczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Teetska
; APPLICANT: Neir, Bruce P.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: FORS-04912
; CURRENT APPLICATION NUMBER: US/09/777,430A
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 647
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-777-430A-72

```

Query Match 98.4%; Score 196.8; DB 9; Length 647;  
Best Local Similarity 54.0%; Pred. No. 1.3e-31;  
Matches 108; Conservative 90; Mismatches 2; Indels 0; Gaps 0;

[illegible]

```

RESULT 9
US-10-875-094-72
; Sequence 72, Application US/10875094
; Publication No. US20050130179A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Szmyczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Tsetska
; APPLICANT: Neri, Bruce P.
; APPLICANT: Prudent, James R.
; APPLICANT: Hall, Jeff G.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: PORS-04912
; CURRENT APPLICATION NUMBER: US/10/875,094
; CURRENT FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 647
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-875-094-72

```

[illegible]

RESULT 10  
US-10-210-120-66  
; Sequence 66, Application US/10210120  
; Publication No. US20030175736A1  
; GENERAL INFORMATION:  
; APPLICANT: Chinnaiyan, Arul M.  
; APPLICANT: Rubin, Mark A.

APPLICANT: Sreekumar, Arun  
TITLE OF INVENTION: Expression Profile of Prostate Cancer  
FILE REFERENCE: UM-07221  
CURRENT APPLICATION NUMBER: US/10/210,120  
CURRENT FILING DATE: 2002-08-01  
PRIOR APPLICATION NUMBER: US 60/309,581  
PRIOR FILING DATE: 2001-08-02  
PRIOR APPLICATION NUMBER: US 60/334,468  
PRIOR FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 66  
LENGTH: 725  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-210-120-66

Query Match 98.4%; Score 196.8; DB 16; Length 725;  
Best Local Similarity 99.0%; Pred. No. 1.3e-31;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 60  
DB 379 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 438  
QY 61 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 120  
DB 439 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 498  
QY 121 TTGATGTTTTAAGTTATCTTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTCGGGA 180  
DB 499 TTGATGTTTTAAGTTATCTTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTCGGGA 558  
QY 181 AATTGCTTTTCCTCTGTAC 200  
DB 559 AATTGCTTTTCCTCTGAAC 578

RESULT 11  
US-10-641-643-1344  
Sequence 1344, Application US/10641643  
Publication No. US20040077003A1  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
Jeffrey J. Seilhamer  
Susan G. Stuart  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
GENE EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSER: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1344:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 725 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g34513  
SEQUENCE DESCRIPTION: SEQ ID NO: 1344 :  
US-10-641-643-1344

Query Match 98.4%; Score 196.8; DB 18; Length 725;  
Best Local Similarity 99.0%; Pred. No. 1.3e-31;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 60  
DB 379 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 438  
QY 61 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 120  
DB 439 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 498  
QY 121 TTGATGTTTTAAGTTATCTTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTCGGGA 180  
DB 499 TTGATGTTTTAAGTTATCTTTTCATGTAAGTACTAGTGTGTTTTAGATACAGAGACTTCGGGA 558  
QY 181 AATTGCTTTTCCTCTGTAC 200  
DB 559 AATTGCTTTTCCTCTGAAC 578

RESULT 12  
US-10-283-975A-58  
Sequence 58, Application US/10283975A  
Publication No. US20040110792A1  
GENERAL INFORMATION:  
APPLICANT: Ortho-Clinical Diagnostics, Inc.  
TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
FILE REFERENCE: CDS 293 PCT  
CURRENT APPLICATION NUMBER: US/10/283,975A  
CURRENT FILING DATE: 2002-10-30  
PRIOR APPLICATION NUMBER: 60/340,938  
PRIOR FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/338,997  
PRIOR FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/340,081  
PRIOR FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/341,012  
PRIOR FILING DATE: 2001-10-30  
NUMBER OF SEQ ID NOS: 900  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 58  
LENGTH: 725  
TYPE: DNA  
ORGANISM: HUMAN  
US-10-283-975A-58

Query Match 98.4%; Score 196.8; DB 19; Length 725;  
Best Local Similarity 99.0%; Pred. No. 1.3e-31;  
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 60  
DB 379 TGCAGCTAACTTATTTCCCTAGCTTCCCGACACACCTGTTTATTATTATAATG 438  
QY 61 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 120  
DB 439 AATTTGTTGTTGATGTGAACATTATGCTTTAAGTAATGTTAAATCTTTATTAAAGTTA 498

```
QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
Db 499 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 558

QY 181 AATTGCTTTTCCTCTTGATC 200
Db 559 AATTGCTTTTCCTCTTGAAC 578

RESULT 13
US-10-764-649-17
; Sequence 17, Application US/10764649
; Publication No. US20040157253A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Haiyan
; APPLICANT: Chen, Hong
; APPLICANT: Barnes, Glenn
; TITLE OF INVENTION: INFLAMMATORY PROTEINS IN THE DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: METABOLIC DISORDERS
; FILE REFERENCE: MPI2003-02SPRNM
; CURRENT FILING DATE: 2004-01-26
; PRIOR FILING DATE: 2004-01-26
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 725
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(353)
US-10-764-649-17

Query Match 98.4%; Score 196.8; DB 19; Length 725;
Best Local Similarity 99.0%; Pred. No. 1.3e-31;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTGTTTTTATTTAATAATG 60
Db 379 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTGTTTTTATTTAATAATG 438

QY 61 AATTTTGTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTTAAGTTA 120
Db 439 AATTTTGTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTTAAGTTA 498

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
Db 499 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 558

QY 181 AATTGCTTTTCCTCTTGATC 200
Db 559 AATTGCTTTTCCTCTTGAAC 578

RESULT 15
US-10-852-335A-46
; Sequence 46, Application US/10852335A
; Publication No. US20050112129A1
; GENERAL INFORMATION:
; APPLICANT: HEIDI S. PHILLIPS
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumors of Glial Origin
; FILE REFERENCE: P5103R1-US
; CURRENT APPLICATION NUMBER: US/10/852,335A
; CURRENT FILING DATE: 2004-05-24
; PRIOR APPLICATION NUMBER: US 60/548,299
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/473,238
; PRIOR FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 46
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-852-335A-46

Query Match 98.4%; Score 196.8; DB 21; Length 725;
Best Local Similarity 99.0%; Pred. No. 1.3e-31;
Matches 198; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTGTTTTTATTTAATAATG 60
Db 379 TGCAGCTAACTTATTTCCCTAGCTTTCCCGACACCTTGTGTTTTTATTTAATAATG 438

QY 61 AATTTTGTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTTAAGTTA 120
Db 439 AATTTTGTGTTGATGTGAACATTATGCTTAAGTAATGTTAATCTTATTTAAGTTA 498

QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
Db 499 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 558

QY 181 AATTGCTTTTCCTCTTGATC 200
Db 559 AATTGCTTTTCCTCTTGAAC 578
```

```
QY 121 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 180
Db 499 TTGATGTTTAAAGTTTATCTTTTCATGCTACTAGTGTGTTTTTAGATACAGAGACTTGGGA 558

QY 181 AATTGCTTTTCCTCTTGATC 200
Db 559 AATTGCTTTTCCTCTTGAAC 578

RESULT 14
US-10-445-790-3
; Sequence 3, Application US/10445790
; Publication No. US20040197305A1
; GENERAL INFORMATION:
; APPLICANT: DeVico, Anthony L.
; APPLICANT: Garzino, Alfredo
; TITLE OF INVENTION: Immuno-Modulating Effects of Chemokines in DNA Vaccination
; FILE REFERENCE: 4115-109 CIP DIV
; CURRENT APPLICATION NUMBER: US/10/445,790
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: PCT/US98/26291
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/591,992
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/186,416
```

Search completed: August 4, 2005, 15:54:12  
Job time : 358.761 secs

---

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 602.915 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-3  
Perfect score: 150  
Sequence: 1 caaagatcacatttagctc.....tcattctaaatgctttttc 150

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150	100.0	150	6	AX343330 Sequence
2	150	100.0	9174	9	AF519531 Homo sapi
3	150	100.0	11793	9	HSY18933 Homo sapien
4	148.4	98.9	73806	2	AC021520 Homo sapi
5	148.4	98.9	147416	9	AC005549 Homo sapi
6	146.4	97.6	2776	9	M37719 Human monoc
7	136.4	90.9	2243	9	HSJBEPR
8	111.6	74.4	5926	6	AX346555 Sequence
9	92	61.3	5926	6	AX346554 Sequence
10	59.6	39.7	3405	4	BOVMCP1X
11	37.4	24.9	109929	10	AC093922 Genomic s
12	37.4	24.9	166277	10	AC093924 Genomic s
13	36.6	24.4	159270	8	AC134230 Oryza sat
14	36.4	24.3	191270	9	AP000776 Homo sapi
15	36.4	24.3	198211	2	AC073756 Mus muscu
16	36	24.0	3590	3	BT003528 Drosophil
17	36	24.0	195769	2	AC147691 Pongo pyg
18	36	24.0	200490	10	AC132388 Mus muscu
19	36	24.0	275076	10	AC093360 Mus muscu

C 20	35.6	23.7	87572	8	AP004135	AP004135 Oryza sat
C 21	35.6	23.7	108091	8	AP004649	AP004649 Oryza sat
C 22	35.6	23.7	143501	9	AC104383	AC104383 Homo sapi
C 23	35.6	23.7	192898	9	AB145968	AB145968 Pan trogl
C 24	35.6	23.7	300050	9	AB100086	AB100086 Pan trogl
C 25	35.4	23.6	43276	9	AC074214	AC074214 Homo sapi
C 26	35.4	23.6	212410	2	AC118804	AC118804 Rattus no
C 27	35.4	23.6	213458	2	AC132561	AC132561 Rattus no
C 28	35.2	23.5	98984	9	AC027220	AC027220 Homo sapi
C 29	35.2	23.5	174347	10	AL928956	AL928956 Mouse DNA
C 30	35.2	23.5	197064	5	BX537347	BX537347 Zebrafish
C 31	35.2	23.5	252997	2	BX927091	BX927091 Danio rer
C 32	35	23.3	68037	9	AY046055	AY046055 Homo sapi
C 33	35	23.3	98876	9	AC009488	AC009488 Homo sapi
C 34	35	23.3	227968	9	AF053356	AF053356 Homo sapi
C 35	34.8	23.2	146254	2	AC141828	AC141828 Apis mell
C 36	34.8	23.2	171297	9	AC096970	AC096970 Homo sapi
C 37	34.8	23.2	190506	2	AC068990	AC068990 Homo sapi
C 38	34.8	23.2	216342	9	AC009132	AC009132 Homo sapi
C 39	34.8	23.2	228361	10	AC117656	AC117656 Mus muscu
C 40	34.6	23.1	12742	1	U67556	U67556 Methanocald
C 41	34.6	23.1	55173	2	AC132816	AC132816 Homo sapi
C 42	34.6	23.1	110000	2	AC121713_2	Continuation (3 of
C 43	34.6	23.1	110000	2	AC121713_3	Continuation (4 of
C 44	34.6	23.1	110000	6	AR271569_03	Continuation (4 of
C 45	34.6	23.1	153171	9	AC147092	AC147092 Pan trogl

ALIGNMENTS

RESULT 1  
AX343330  
LOCUS AX343330  
DEFINITION Sequence 3 from Patent EP1170372.  
ACCESSION AX343330  
VERSION AX343330.1 GI:18491680  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Roesl,F., Soto,U., Coy,J., Finzer,P., Delius,H., Poustka,A., zur Hausen,H. and Patzelt,A.  
TITLE Regulatory sequences of the human mcp-1 gene  
JOURNAL Patent: EP 1170372-A 3 09-JAN-2002;  
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OFFENTLICHEN RECHTS (DE)  
FEATURES  
Location/Qualifiers  
1..150  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 150; DB 6; Length 150;  
Best Local Similarity 100.0%; Pred.No. 3.1e-31;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CAAAGATCACATTTCTAGCTCTGAGGTATAGGAGGACACCTGGGATTTTAAATGAGCTCTTTT 60  
Db 1 CAAAGATCACATTTCTAGCTCTGAGGTATAGGAGGACACCTGGGATTTTAAATGAGCTCTTTT 60  
Qy 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCTCTCATGACTCTTTTCTGCTCTTAAAGATCAGA 120  
Db 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCTCTCATGACTCTTTTCTGCTCTTAAAGATCAGA 120  
Qy 121 ATAATCCAGTTTCATCTTAAATGCTTTTTC 150  
Db 121 ATAATCCAGTTTCATCTTAAATGCTTTTTC 150

us-09-899-276c-3.rge

Tue Aug 9 17:01:25 2005

```

RESULT 2
AF519531      9174 bp   DNA      linear   PRI 09-JUL-2002
LOCUS        Homo sapiens small inducible cytokine A2 (monocyte chemotactic
DEFINITION
ACCESSION   AF519531
VERSION     AF519531
KEYWORDS    AF519531.1 GI:21435976
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 9174)
AUTHORS    Rieder, M.J., Armel, T.Z., Carrington, D.P., Ozuna, M., Kuldane, S.A.,
Rajkumar, N., Toth, E.J., Yi, Q. and Nickerson, D.A.
TITLE      Direct Submission
JOURNAL    Submitted (06-JUN-2002) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
COMMENT    To cite this work please use: SeattleSNPs. NHLBI HL66682 Program
for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source           1..9174
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     variation        107
                     /frequency="0.17"
                     /replace="g"
     variation        291
                     /frequency="0.03"
                     /replace="g"
     variation        305
                     /frequency="0.15"
                     /replace="t"
     variation        334
                     /frequency="0.01"
                     /replace="g"
     variation        604
                     /frequency="0.16"
                     /replace="t"
     variation        609
                     /frequency="0.02"
                     /replace="a"
     repeat_region    614..720
                     /rpt_family="MIR"
                     /rpt_type="dispersed"
     variation        678
                     /frequency="0.14"
                     /replace="g"
     variation        1051
                     /frequency="0.01"
                     /replace="a"
     variation        1054
                     /frequency="0.01"
                     /replace="t"
     variation        1072
                     /frequency="0.03"
                     /replace="t"
     variation        1280
                     /frequency="0.23"
                     /replace="a"
     variation        1537
                     /frequency="0.15"
                     /replace="g"
     variation        1548
                     /frequency="0.08"
                     /replace="a"
     variation        1563
                     /frequency="0.04"
                     /replace="g"
     variation        1707
                     /frequency="0.02"
repeat_region        /replace="c"
1718..1830
/rpt_family="MIR"
/rpt_type="dispersed"
1729
/frequency="0.01"
/replace="c"
1980
/frequency="0.13"
/replace="t"
1984
/frequency="0.02"
/replace="g"
2307
/frequency="0.04"
/replace="g"
2391
/frequency="0.01"
/replace="a"
2569
/frequency="0.08"
/replace="t"
3001
/frequency="0.03"
/replace="a"
3190
/frequency="0.14"
/replace="c"
3228..3445
/rpt_family="MIR"
/rpt_type="dispersed"
3719
/frequency="0.01"
/replace="a"
3756
/frequency="0.25"
/replace="c"
4065..5981
/gene="SCYA2"
join(4065..4193,4990..5107,5490..5981)
/gene="SCYA2"
/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
join(4118..4193,4990..5107,5490..5595)
/gene="SCYA2"
/codon_start=1
/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
/protein_id="AAMS4046.1"
/db_xref="GI:21435977"
/translation="MKVSAALLCLLLIAATFPGQLAQPDAINAPVTCYNTNRKIS
VORLASVRRITSSKCPKEAVIFKTIKAIKADPKQKWVDSMDHLDKQTPTKT"
4296..4409
/rpt_family="MIR"
/rpt_type="dispersed"
4359..4462
/gene="SCYA2"
/Note="Region not scanned for variation"
4748
/gene="SCYA2"
/frequency="0.26"
/replace="a"
4777..4814
/rpt_family="MIR"
/rpt_type="dispersed"
4881
/gene="SCYA2"
/frequency="0.14"
/replace="g"
5018
/gene="SCYA2"
/frequency="0.47"
/replace="c"

```

variation 5132  
/gene="SCYA2"  
/frequency="0.01"  
/replace="t"  
variation 5660  
/gene="SCYA2"  
/frequency="0.18"  
/replace="t"  
variation 6184  
/frequency="0.01"  
/replace="c"  
variation 6518  
/frequency="0.01"  
/replace="t"  
variation 6531  
/frequency="0.05"  
/replace="t"  
repeat\_region 6837..6916  
/rpt\_family="L2"  
/rpt\_type=dispersed  
variation 7436  
/frequency="0.07"  
/replace="t"  
variation 7605  
/frequency="0.03"  
/replace="a"  
variation 7614  
/frequency="0.01"  
/replace="a"  
variation 7659  
/frequency="0.03"  
/replace="a"  
variation 7710  
/frequency="0.05"  
/replace="a"  
variation 7843  
/frequency="0.31"  
/replace="c"  
variation 8072  
/frequency="0.01"  
/replace="t"  
repeat\_region 8118..8271  
/rpt\_family="MaLR"  
/rpt\_type=dispersed  
repeat\_region 8272..8734  
/rpt\_family="MaLR"  
/rpt\_type=dispersed  
variation 8597  
/frequency="0.07"  
/replace="c"  
repeat\_region 8735..9033  
/rpt\_family="Alu"  
/rpt\_type=dispersed  
variation 8775  
/frequency="0.07"  
/replace="a"

Query Match 100.0%; Score 150; DB 9; Length 9174;  
Best Local Similarity 100.0%; Pred. NO. 1.4e-31;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGATCATCTTAGCTCTGAGGTATAGGACGAGCACTGGGATTTAATGAGCTCTTT 60  
|||||  
Db 4821 CAAAGATCATCTTAGCTCTGAGGTATAGGACGAGCACTGGGATTTAATGAGCTCTTT 4880  
|||||  
QY 61 CTCCTCTCTGCTGCTTTTGGCTTTTCTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 120  
|||||  
Db 4881 CTCCTCTCTGCTGCTTTTGGCTTTTCTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 4940  
|||||  
QY 121 ATAATCCAGTTTCATCTTAATGCTTTTTC 150  
|||||  
Db 4941 ATAATCCAGTTTCATCTTAATGCTTTTTC 4970  
|||||

RESULT 3  
LOCUS HSY18933 11793 bp DNA linear PRI 18-OCT-2000  
DEFINITION Homo sapiens MCP-1 gene and enhancer region.  
ACCESSION Y18933  
VERSION Y18933.1 GI:10933860  
KEYWORDS MCP-1 gene; monocyte chemoattractant protein-1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Finzer,P., Soto,U., Delius,H., Patzelt,A., Coy,J.F., Poustka,A.,  
zur Hausen,H. and Roel,F.  
TITLE Differential transcriptional regulation of the  
monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and  
non-tumorigenic HPV 18 positive cells: the role of the chromatin  
structure and AP-1 composition  
JOURNAL Oncogene 19 (29), 3235-3244 (2000)  
MEDLINE 20374005  
PubMed 10918580  
REFERENCE 2 (bases 1 to 11793)  
AUTHORS Roesl,F.  
TITLE Direct Submission  
JOURNAL Submitted (15-APR-1999) F. Roesl, Applied Tumor Virology, DKFZ  
(German Cancer Research Institute), Im Neuenheimer Feld 506, 69120  
Heidelberg, FRG  
FEATURES  
source  
1..11793  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
4906..9073  
/gene="MCP-1"  
4906..4915  
/gene="MCP-1"  
4935..4944  
/gene="MCP-1"  
7513..7517  
/gene="MCP-1"  
Join(<7550..7685,8468..8585,8968..>9073)  
/gene="MCP-1"  
<7550..7685  
/gene="MCP-1"  
/number=1  
Join(7610..7685,8468..8585,8968..9073)  
/gene="MCP-1"  
/codon\_start=1  
/product="monocyte chemoattractant protein-1"  
/protein\_id="CAC14049.1"  
/db\_xref="GI:10933861"  
/db\_xref="GOA:P13500"  
/db\_xref="UniProt/Swiss-Prot:P13500"  
/translation="MKVSAALLCLLLIAATFIPQGLAOPDAINAPVTCYNTNRKIS  
VORLASVERITSSKCPKEAVIFKTIKAIKADPKQWQSDMDHLDKQTQPKT"  
7686..8467  
/gene="MCP-1"  
/number=1  
8468..8585  
/gene="MCP-1"  
/number=2  
8586..8967  
/gene="MCP-1"  
/number=2  
8968..>9073  
/gene="MCP-1"  
/number=3  
ORIGIN  
Query Match 100.0%; Score 150; DB 9; Length 11793;  
Best Local Similarity 100.0%; Pred. No. 1.3e-31;

Tue Aug 9 17:01:25 2005

Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGATCACATCTAGCTCTGAGGTATAGCAGAGCACTGGGATTAATGAGCTCTTT 60  
 |||||  
 Db 8299 CAAGATCACATCTAGCTCTGAGGTATAGCAGAGCACTGGGATTAATGAGCTCTTT 8358  
 |||||

QY 61 CTCTCTCTGCTGCTGCTTTTTCCTTCATGACTCTTTCTGCTCTTAAGATCAGA 120  
 |||||  
 Db 8359 CTCTCTCTGCTGCTGCTTTTTCCTTCATGACTCTTTCTGCTCTTAAGATCAGA 8418  
 |||||

QY 121 ATAATCCAGTTCATCTAAATGCTTTTTC 150  
 |||||  
 Db 8419 ATAATCCAGTTCATCTAAATGCTTTTTC 8448  
 |||||

RESULT 4  
 AC021520/c 73806 bp DNA linear HTG 13-JUL-2000  
 LOCUS Homo sapiens chromosome 17 clone RP11-329H16 map 17, LOW-PASS  
 DEFINITION SEQUENCE SAMPLING.

AC021520.2 GI:9148483  
 VERSION HTG; HTGS PHASE0.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 73806)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 17, clone RP11-329H16  
 Unpublished  
 2 (bases 1 to 73806)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
 Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,  
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,  
 DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestock,J.,  
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,  
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,  
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
 Zimmer,A. and Zody,M.

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 13, 2000 this sequence version replaced GI:6705580.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L5499  
 Center clone name: 329\_H16  
 -----

\* NOTE: This record contains 76 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone

\* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 862: contig of 862 bp in length  
 962: gap of 100 bp  
 1850: contig of 888 bp in length  
 1851: gap of 100 bp  
 1850: contig of 867 bp in length  
 1951: gap of 100 bp  
 2817: contig of 871 bp in length  
 2818: gap of 100 bp  
 3788: contig of 872 bp in length  
 3789: gap of 100 bp  
 4760: contig of 865 bp in length  
 4761: gap of 100 bp  
 5725: contig of 865 bp in length  
 5726: gap of 100 bp  
 5825: contig of 895 bp in length  
 5826: gap of 100 bp  
 6721: gap of 100 bp  
 6721: contig of 860 bp in length  
 7681: gap of 100 bp  
 7681: contig of 886 bp in length  
 8667: gap of 100 bp  
 8667: contig of 871 bp in length  
 9637: gap of 100 bp  
 9638: contig of 877 bp in length  
 10614: gap of 100 bp  
 10614: contig of 893 bp in length  
 11607: gap of 100 bp  
 11607: contig of 876 bp in length  
 12583: gap of 100 bp  
 12583: contig of 887 bp in length  
 12684: gap of 100 bp  
 12684: contig of 890 bp in length  
 13671: gap of 100 bp  
 13671: contig of 876 bp in length  
 14561: gap of 100 bp  
 14561: contig of 885 bp in length  
 15537: gap of 100 bp  
 15537: contig of 897 bp in length  
 16622: gap of 100 bp  
 16622: contig of 898 bp in length  
 17618: gap of 100 bp  
 17618: contig of 884 bp in length  
 18516: gap of 100 bp  
 18516: contig of 884 bp in length  
 19501: gap of 100 bp  
 19501: contig of 855 bp in length  
 20455: gap of 100 bp  
 20455: contig of 886 bp in length  
 21441: gap of 100 bp  
 21441: contig of 877 bp in length  
 22418: gap of 100 bp  
 22418: contig of 858 bp in length  
 23376: gap of 100 bp  
 23376: contig of 866 bp in length  
 24342: gap of 100 bp  
 24342: contig of 791 bp in length  
 25233: gap of 100 bp  
 25233: contig of 890 bp in length  
 26223: gap of 100 bp  
 26223: contig of 871 bp in length  
 27194: gap of 100 bp  
 27194: contig of 868 bp in length  
 28162: gap of 100 bp  
 28162: contig of 886 bp in length  
 29248: gap of 100 bp  
 29248: contig of 906 bp in length  
 30154: gap of 100 bp  
 30154: contig of 870 bp in length  
 31124: gap of 100 bp  
 31124: contig of 881 bp in length  
 32105: gap of 100 bp  
 32105: contig of 888 bp in length  
 33093: gap of 100 bp  
 33093: contig of 885 bp in length  
 34178: gap of 100 bp  
 34178: contig of 885 bp in length

\* 34179 34826: contig of 648 bp in length  
\* 34827 34926: gap of 100 bp  
\* 35771 35771: contig of 845 bp in length  
\* 35772 35871: gap of 100 bp  
\* 35872 36751: contig of 880 bp in length  
\* 36752 36851: gap of 100 bp  
\* 36852 37741: contig of 890 bp in length  
\* 37742 37841: gap of 100 bp  
\* 37842 38705: contig of 864 bp in length  
\* 38706 38805: gap of 100 bp  
\* 38806 39678: contig of 873 bp in length  
\* 39679 39778: gap of 100 bp  
\* 39779 40673: contig of 895 bp in length  
\* 40674 40773: gap of 100 bp  
\* 40774 41671: contig of 898 bp in length  
\* 41672 41771: gap of 100 bp  
\* 41772 42631: contig of 860 bp in length  
\* 42632 42731: gap of 100 bp  
\* 42732 43606: contig of 875 bp in length  
\* 43607 43706: gap of 100 bp  
\* 43707 44539: contig of 833 bp in length  
\* 44540 44639: gap of 100 bp  
\* 44640 45524: contig of 885 bp in length  
\* 45525 45624: gap of 100 bp  
\* 45626 46512: contig of 888 bp in length  
\* 46513 46612: gap of 100 bp  
\* 46613 47479: contig of 867 bp in length  
\* 47480 47579: gap of 100 bp  
\* 47580 48478: contig of 899 bp in length  
\* 48479 48578: gap of 100 bp  
\* 48579 49448: contig of 870 bp in length  
\* 49449 49548: gap of 100 bp  
\* 49549 50430: contig of 882 bp in length  
\* 50431 50530: gap of 100 bp  
\* 50531 51430: contig of 900 bp in length  
\* 51431 51530: gap of 100 bp  
\* 51531 52409: contig of 879 bp in length  
\* 52410 52509: gap of 100 bp  
\* 52510 53385: contig of 876 bp in length  
\* 53386 53485: gap of 100 bp  
\* 53486 54376: contig of 891 bp in length  
\* 54377 54476: gap of 100 bp  
\* 54477 55333: contig of 857 bp in length  
\* 55334 55433: gap of 100 bp  
\* 55434 56289: contig of 856 bp in length  
\* 56290 56389: gap of 100 bp  
\* 56390 57282: contig of 893 bp in length  
\* 57283 57382: gap of 100 bp  
\* 57383 58278: contig of 896 bp in length  
\* 58279 59254: contig of 876 bp in length  
\* 59255 59354: gap of 100 bp  
\* 59355 60220: contig of 866 bp in length  
\* 60221 60320: gap of 100 bp  
\* 60321 61212: contig of 892 bp in length  
\* 61213 62196: contig of 884 bp in length  
\* 62197 62296: gap of 100 bp  
\* 62297 63167: contig of 871 bp in length  
\* 63168 63267: gap of 100 bp  
\* 63268 64169: contig of 902 bp in length  
\* 64170 64269: gap of 100 bp  
\* 64270 65168: contig of 899 bp in length  
\* 65169 65268: gap of 100 bp  
\* 65269 66137: contig of 869 bp in length  
\* 66138 66237: gap of 100 bp  
\* 66238 67123: contig of 886 bp in length  
\* 67124 68104: contig of 881 bp in length  
\* 68105 68204: gap of 100 bp  
\* 68205 69094: contig of 890 bp in length

Query Match 98.9%; Score 148.4; DB 2; Length 73806;

Best Local Similarity 99.3%; Pred. No. 2.6e-31;  
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CAAAGATCACATTTCTAGCTCTGAGGTATAGGACAGAGCACTGGGATTTTAATGAGCTCTTT 60  
DB 57929 CAAAGATCACATTTCTAGCTCTGAGGTATAGGACAGAGCACTGGGATTTTAATGAGCTCTTT 57870  
QY 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 120  
DB 57869 CTCCTCTCATGCTGCTTTTCTTTTCTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 57810  
QY 121 ATAAATCCAGTTTCATCTCTAAATGCTTTTTC 150  
DB 57809 ATAAATCCAGTTTCATCTCTAAATGCTTTTTC 57780  
RESULT 5  
AC005549/c 147416 bp DNA linear PRI 22-SEP-1998  
LOCUS Homo sapiens chromosome 17, clone hRPK.215\_E13, complete sequence.  
DEFINITION AC005549  
ACCESSION AC005549  
VERSION AC005549.1 GI:3598724  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 147416)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
REFERENCE  
1 (bases 1 to 147416)  
Homo sapiens chromosome 17, clone hRPK.215\_E13  
2 (bases 1 to 147416)  
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J.,  
Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M.,  
Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E.,  
Devon,K., Dewar,K., Donelan,L., Ferreira,P., Fitzhugh,W.,  
Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G.,  
Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C.,  
Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J.,  
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,  
Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P.,  
Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,  
Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,  
Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I.,  
Vassiliev,H., Vo.A., Wagner,A., Wheeler,J., Wu.Y., Wyman,D.,  
Ye,W.J., Zhao,J. and Zody,W.  
Direct Submission  
Submitted (27-AUG-1998) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 147416)  
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J.,  
Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M.,  
Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E.,  
Devon,K., Dewar,K., Donelan,L., Ferreira,P., Fitzhugh,W.,  
Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G.,  
Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C.,  
Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J.,  
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,  
Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P.,  
Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,  
Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,  
Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I.,  
Vassiliev,H., Vo.A., Wagner,A., Wheeler,J., Wu.Y., Wyman,D.,  
Ye,W.J., Zhao,J. and Zody,W.  
Direct Submission  
Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 147416)  
Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,

Tue Aug 9 17:01:25 2005

Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E., Dewar, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W., Grant, G., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraltgery, K., Grant, G., Hagos, B., Heaford, A., Herena, L., Horton, L., Howland, J.C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Melldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahr, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolasky, N., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W.J., Zhao, J. and Zody, M.

Direct Submission  
Submitted (22-SEP-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 13, 1998 this sequence version replaced gi:3581743.  
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Only the first 147416 bases of this clone are being submitted.  
Bases 145417-155040 overlap accession number AC004147 (WICGR project L228). The first 2Kb of the overlapping region are submitted to confirm overlap.

# FEATURES

## source

1. .147416  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="hBPK.215.E.13"  
/clone\_lib="RPC1-11 human BAC library"

2. .69  
/rpt\_family="BURI1"

repeat\_region  
/rpt\_family="BURI1"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="L1ME2"

repeat\_region  
/rpt\_family="MER5A"  
complement(11031..11072)  
/rpt\_family="MIR"  
11567..11836  
/rpt\_family="L12"  
12501..12981  
/rpt\_family="MER9"  
complement(12983..13257)  
/rpt\_family="AluJb"  
15248..15364  
/rpt\_family="MIR"  
complement(16483..16653)  
/rpt\_family="MIR"  
16830..16949  
/rpt\_family="MER5A"  
17203..17380  
/rpt\_family="LTR16A"  
17393..18306  
/rpt\_family="HERVL"  
18315..24307  
/rpt\_family="L1PA2"  
24309..24335  
/rpt\_family="AT\_rich"  
24513..25245  
/rpt\_family="HERVL"  
25613..25758  
/rpt\_family="HERVL"  
26135..26228  
/rpt\_family="LTR16A"  
complement(27814..27950)  
/rpt\_family="MER94"  
28930..28968  
/rpt\_family="AT\_rich"  
29960..30341  
/rpt\_family="MLT11"  
32364..32581  
/rpt\_family="MIR"  
32601..32902  
/rpt\_family="AluSx"  
complement(33593..39359)  
/rpt\_family="L1PA2"  
39362..39414  
/rpt\_family="A-rich"  
complement(39920..40047)  
/rpt\_family="MIR"  
40112..40151  
/rpt\_family="AT\_rich"  
complement(40152..40442)  
/rpt\_family="AluJo"  
complement(40810..40898)  
/rpt\_family="MLT1J"  
41293..41469  
/rpt\_family="L1PA6"  
complement(41846..41956)  
/rpt\_family="purine-rich"  
complement(42486..42569)  
/rpt\_family="MLT11"  
42975..43098  
/rpt\_family="MER5A"  
43554..43983  
/rpt\_family="LTR24"  
44781..44811  
/rpt\_family="AT\_rich"  
46184..46296  
/rpt\_family="MIR"  
complement(47075..47140)  
/rpt\_family="MLT1G"  
complement(47266..47307)  
/rpt\_family="(CA)n"  
complement(47363..47442)  
/rpt\_family="MLT1G"  
complement(48165..48209)

repeat\_region

```
Query Match      98.9%; Score 148.4; DB 9; Length 147416;
Best Local Similarity 99.3%; Pred. No. 2.3e-31;
Matches 149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAAGATCACAATTCTAGCTCTGAGGTATAGGCGAAGCACTGGGATTTAATGAGCTCTTT 60
Db 60714 CAAAGATCACAATTCTAGCTCTGAGGTATAGGCGAAGCACTGGGATTTAATGAGCTCTTT 60655

QY 61 CTCTTCTCCGCTGCTGCTTTTCTCTCTATGACTCTTTTCTGCTCTTAAGATCAGA 120
Db 60654 GTCTTCTCCGCTGCTGCTTTTCTCTCTATGACTCTTTTCTGCTCTTAAGATCAGA 120

QY 121 ATAATCCAGTTCATCTCTAAATGCTTTTC 150
Db 60594 ATAATCCAGTTCATCTCTAAATGCTTTTC 60565

RESULT 6
HUMMCHEMP
LOCUS Human monocyte chemotactic protein gene, complete cds.
ACCESSION M37719
VERSION M37719.1 GI:187447
KEYWORDS monocyte chemotactic protein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2776).
AUTHORS Shyy.Y.J., Li,Y.S. and Kolattukudy,P.E.
TITLE Structure of human monocyte chemotactic protein gene and its
regulation by TPA
JOURNAL Biochem. Biophys. Res. Commun. 169 (2), 346-351 (1990)
MEDLINE 90290466
PUBMED 2357211
COMMENT Original source text: Human DNA.
FEATURES
    source
        1..2776
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
        598..2080
            /genes="SCYA2"
            join(598..673,1472..1589,1975..2080)
            /genes="SCYA2"
            /note="monocyte chemotactic protein"
            /codon_start=1
            /protein_id="AAAI8102.1"
            /db_xref="GI:487124"
            /translation="MKVSAALICLLIAATFIPQGLAOPDAINAPVTCYNTFRKIS
            VORLASVRRITSSKCPKEAVIFKTIIVAKEICADPKQKWQDSMDHLDKQTQPKT"
            <598..673
            /genes="SCYA2"
            /note="monocyte chemotactic protein"
            /number=1
        598..673
            /genes="SCYA2"
            /note="monocyte chemotactic protein"
            /number=1
        674..1471
            /genes="SCYA2"
            /note="monocyte chemotactic protein intron A"
            1472..1589
            /genes="SCYA2"
            /note="monocyte chemotactic protein"
            /number=2
            1590..1974
            /genes="SCYA2"
            /note="monocyte chemotactic protein intron B"
            1975..2080
            /genes="SCYA2"
            /note="monocyte chemotactic protein"
            /number=3

exon
    1975..2080
        /gene="SCYA2"
        /note="monocyte chemotactic protein"
        /number=3

ORIGIN
Query Match      97.6%; Score 146.4; DB 9; Length 2776;
Best Local Similarity 99.3%; Pred. No. 1.8e-30;
Matches 147; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAAGATCACAATTCTAGCTCTGAGGTATAGGCGAAGCACTGGGATTTAATGAGCTCTTT 60
Db 1304 CAAAGATCACAATTCTAGCTCTGAGGTATAGGCGAAGCACTGGGATTTAATGAGCTCTTT 1363

QY 61 CTCTTCTCTGCTGCTGCTTTTCTCTCTATGACTCTTTTCTGCTCTTAAGATCAGA 120
Db 1364 CTCTTCTCTGCTGCTGCTTTTCTCTCTATGACTCTTTTCTGCTCTTAAGATCAGA 120

QY 121 ATAATCCAGTTCATCTCTAAATGCTTTT 148
Db 1424 ATAATCCAGTTCATCTCTAAATGCTTTT 1451

RESULT 7
HSJEP
LOCUS H.SJEP 2243 bp DNA linear PRI 03-NOV-1992
DEFINITION H.sapiens gene for JE protein, exons 3 and 4.
ACCESSION X60001
VERSION X60001.1 GI:34001
KEYWORDS JE protein; secretory protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2243)
AUTHORS Schwarz,E., Fan,E., Kallin,B., Sorrentino,V. and Bloom,B.
TITLE The IFN-gamma inducible cDNA gamma.1 is an incompletely spliced JE
cDNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2243)
AUTHORS Schwarz,E.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-1991) E. Schwarz, Albert Einstein College of
Medicine, Dept of Microbiology & Immunology, Forchheimer Bldg Room
411, 1300 Morris Park Avenue, Bronx NY 10461, USA

FEATURES
    source
        1..2243
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /tissue_type="liver"
            /clone_lib="Okayama-Berg plasmid"
            /dev_stage="fetal"
            <1..612
            /number=1
        613..1591
            /gene="gamma 1,hje"
            613..730
            /gene="gamma 1,hje"
            /product="JE protein"
            /number=2
            731..1112
            /gene="gamma 1,hje"
            /number=2
            1113..1591
            /gene="gamma 1,hje"
            /product="JE protein"
            /number=3

intron
    613..1591
        /gene="gamma 1,hje"
        /number=2
    613..730
        /gene="gamma 1,hje"
        /product="JE protein"
        /number=2
    731..1112
        /gene="gamma 1,hje"
        /number=2
    1113..1591
        /gene="gamma 1,hje"
        /product="JE protein"
        /number=3

exon
    613..1591
        /gene="gamma 1,hje"
        /product="JE protein"
        /number=3

ORIGIN
Query Match      90.9%; Score 136.4; DB 9; Length 2243;
Best Local Similarity 98.7%; Pred. No. 1.2e-27;
Matches 148; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```





<b>FEATURES</b>					
source	Location/Qualifiers				
	1. .166277				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
	/chromosome="17"				
	/clone="RP23-239M9"				
	/clone_lib="RPCI-23"				
misc_feature	8201..8252				
	/note="We believe the assembly to be correct. The sequence is covered by a single high quality subclone which was sequenced with dye terminator chemistry."				
<b>ORIGIN</b>					
Query Match	24.9%; Score 37.4; DB 10; Length 166277;				
Best Local Similarity	60.2%; Pred. No. 2.3;				
Matches 62; Conservative	0; Mismatches 41; Indels 0; Gaps 0;				
QY	22 GAGGTATAGGCAGNAGCACTGGGATTTAATCAGCTCTTTCCTCTCTCTCCCTCGCCCTTTT 81				
Db	37149 GAGGAGAGGAGAGAGGAGGATTATTCATCTTTTCTTGACCCATCCAGCTATT 37090				
QY	82 GCTTTTTCTCATGACTCTTTTCTGCTTTTAAGATCAGATAA 124				
Db	37089 TCTTTTCTGCTTCCTTTCTTCTCTCTTAAGCTGACATTCA 37047				
<b>RESULT 13</b>					
AC134230	159270 bp DNA linear PLN 11-MAR-2003				
LOCUS	Oryza sativa (japonica cultivar-group) chromosome 3 clone				
DEFINITION	OSJNBa0005J15, complete sequence.				
ACCESSION	AC134230				
VERSION	AC134230.2 GI:28913002				
KEYWORDS	HTG.				
SOURCE	Oryza sativa (japonica cultivar-group)				
ORGANISM	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.				
REFERENCE	1 (bases 1 to 159270) Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K. Rice Genomic Sequence Unpublished 2 (bases 1 to 159270) Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J. and Collura,K. Direct Submission Submitted (24-SEP-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA 3 (bases 1 to 159270) Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K. Direct Submission Submitted (11-MAR-2003) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA On Mar 11, 2003 this sequence version replaced gi:23306132. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone; and the assembly was confirmed by restriction digest. The nucleotide sequence of this BAC clone was generated by combining Syngenta, Monsanto and Arizona Genomics Institute sequencing data. There is only Syngenta reads from base 55503-57821. From base 123135-123143 there is an A run where the Syngenta reads express 10 A's and the AGI reads say 9. There is a run of about 15 C's from base 132783-132796 that has an A pop up in a couple of reads. From base 139551-139618 there are 21-30 TA tandem repeats. There are Bacterial Transposons at the				
TITLE					
JOURNAL					
COMMENT					

```
Db 154300 ACCTAGCATCTTCTCTCAGGCGCTTTTGTGTAATAATATCAATGTAATAGATTATT 154241
QY 134 TCCTAAATGCTTT 147
Db 154240 TGTGAATATAATAT 154227

RESULT 15
AC073756/c
LOCUS 198211 bp DNA linear HTG 18-JUL-2000
DEFINITION Mus musculus clone RP23-336A14, WORKING DRAFT SEQUENCE, 24 ordered
pieces.
ACCESSION AC073756
VERSION AC073756.2 GI:9256783
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 198211)
JOURNAL DOE Joint Genome Institute.
REFERENCE 2 (bases 1 to 198211)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
COMMENT Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:8810373.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1865910
Center clone name: RPCI-23_336A14
-----
Summary Statistics
Consensus quality: 181038 bases at least Q40
Consensus quality: 192296 bases at least Q30
Consensus quality: 194577 bases at least Q20
Estimated insert size: 194000; agarose-fp estimation
Estimated insert size: 197111; sum-of-contigs estimation
Quality coverage: 6.22 in Q20 bases; agarose-fp estimation
Quality coverage: 6.12 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 19138: contig of 19138 bp in length
* 19139 19238: gap of unknown length
* 19239 28743: contig of 9505 bp in length
* 28744 28843: gap of unknown length
* 28844 30952: contig of 2109 bp in length
* 30953 31052: gap of unknown length
* 31053 52408: contig of 21356 bp in length
* 52409 52508: gap of unknown length
* 52509 57315: contig of 4807 bp in length
* 57316 72225: gap of unknown length
* 72226 72325: gap of unknown length
* 72326 77324: contig of 4999 bp in length
* 77325 77424: gap of unknown length
* 77425 87512: contig of 10088 bp in length
* 87513 95934: gap of unknown length
* 87613 95934: contig of 8322 bp in length

95935 96034: gap of unknown length
* 96035 100602: contig of 4568 bp in length
* 100603 100702: gap of unknown length
* 100703 105315: contig of 4613 bp in length
* 105316 105416: gap of unknown length
* 105417 106516: contig of 1101 bp in length
* 106517 106616: gap of unknown length
* 106617 127601: contig of 20985 bp in length
* 127602 127701: gap of unknown length
* 127702 132458: contig of 4757 bp in length
* 132459 132558: gap of unknown length
* 132559 150098: contig of 17540 bp in length
* 150099 150198: gap of unknown length
* 150199 156969: contig of 6771 bp in length
* 156970 157069: gap of unknown length
* 157070 159491: contig of 2422 bp in length
* 159492 159591: gap of unknown length
* 159592 173169: contig of 13578 bp in length
* 173170 173269: gap of unknown length
* 173270 176743: contig of 3474 bp in length
* 176744 176843: gap of unknown length
* 176844 185337: contig of 8634 bp in length
* 185338 185637: gap of unknown length
* 185638 186938: contig of 1301 bp in length
* 186939 187038: gap of unknown length
* 187039 190139: contig of 3101 bp in length
* 190140 190239: gap of unknown length
* 190240 193371: contig of 3132 bp in length
* 193372 193471: gap of unknown length
* 193472 198211: contig of 4740 bp in length.

FEATURES
Location/Qualifiers
1..198211
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-336A14"
/clone_lib="RPCI mouse BAC library 23"

ORIGIN

Query Match 24.3%; Score 36.4; DB 2; Length 198211;
Best Local Similarity 58.2%; Pred. No. 4.2;
Matches 64; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 33 AGAAGCACATGGGATTAATAGACTCTTTCTCTCTCCCTGCGCTTTTGTCTTTTCCCTC 92
Db 156512 AGAAGCAGGTGATCCCTTGAGCCACTTTTCAGCAACAGGCTGTTTCTTGTGACATCC 156453
QY 93 ATGACTCTTTTCTGCTCTTAAGATCAGATAATCCAGTTTCATCTCTAAAT 142
Db 156452 AGGACTCTTTGATCCACTTCAGATCAAGTCACATCCATTTTCTCTAAAT 156403

Search completed: August 4, 2005, 11:16:45
Job time : 608.915 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 3, 2005, 15:09:47 ; Search time 99.413 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-3  
Perfect score: 150  
Sequence: 1 caaagatcacatttagctc.....tcactcaaaagtcttttc 150

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150	100.0	150	6	ADH13940 Human mon
2	150	100.0	724	8	ABX63839 Human CDN
3	150	100.0	11793	12	ADO003803 Human Ccl
4	149.6	99.7	750	8	ACF64399 Human MCP
5	146	97.3	2775	6	ABK47979 Human sma
6	111.6	74.4	5926	6	ABL33653 Human imm
7	92	61.3	5926	6	ABL33652 Human imm
8	58	38.7	840	6	ABK47980 Human sma
9	53	35.3	201	10	ADC98306 SCY22 gen
10	35	23.3	227968	6	ABK83497 Human cdn
11	35	23.3	227968	12	ADQ18538
12	34.6	23.1	110000	2	AAV21209_03
13	34	22.7	2067	4	AAI64628 Mouse MIT
14	34	22.7	2125	4	AAI64631 Mouse MIT
15	34	22.7	2196	4	AAI64627 Mouse MIT
16	33.6	22.4	81684	13	ABD33502 Murine ca
17	33.4	22.3	2631	12	ADNA1804
18	33.2	22.1	285	5	AA883745 Novel hum
19	33.2	22.1	734	5	AA883748 DNA encod
20	33.2	22.1	2275	12	ADQ83497 Human tum

C 21	33.2	22.1	2275	12	ADQ86732	Adq86732 Human tum
C 22	33.2	22.1	2275	13	ADQ84252	Adq84252 Human tum
C 23	33.2	22.1	2275	13	ADQ85632	Adq85632 Human tum
C 24	32.8	21.9	467	9	ACH26101	Ach26101 Human adu
C 25	32.6	21.7	60604	12	ADO48003	Ado48003 Human HIP
C 26	32.4	21.6	37160	6	ABK84372	Abk84372 Human cdn
C 27	32	21.3	407	6	ABN96598	Abn96598 Gene #309
C 28	32	21.3	6879	13	ADR88468	Adr88468 Mouse cdn
C 29	32	21.3	10708	6	ABL68626	AbL68626 Kidney ca
C 30	32	21.3	38059	4	AAF54018	Aaf54018 Human fac
C 31	32	21.3	38059	6	ABN95627	Abn95627 Gene #212
C 32	32	21.3	38059	10	ADD71098	Add71098 Human coa
C 33	31.8	21.2	3747	10	ADF81731	Adf81731 Leukaemia
C 34	31.8	21.2	6184	4	ABL26050	AbL26050 Drosophel
C 35	31.8	21.2	8483	6	ABD36766	Abd36766 Human mel
C 36	31.8	21.2	122557	12	ADH76849	Adh76849 Melanin-c
C 37	31.6	21.1	2734	10	ADJ92180	Adj92180 Human hai
C 38	31.6	21.1	5020	5	AA83075	Aa83075 DNA encod
C 39	31.6	21.1	5039	5	AA83056	Aa83056 DNA encod
C 40	31.6	21.1	5071	5	AA83061	Aa83061 DNA encod
C 41	31.6	21.1	5423	5	AA83062	Aa83062 DNA encod
C 42	31.6	21.1	5423	6	ABL68483	AbL68483 Kidney ca
C 43	31.6	21.1	5655	5	AA83055	Aa83055 DNA encod
C 44	31.6	21.1	5791	5	AA83065	Aa83065 DNA encod
C 45	31.6	21.1	5800	5	AA83054	Aa83054 DNA encod

## ALIGNMENTS

RESULT 1  
ADH13940  
ID ADH13940 standard; DNA; 150 BP.  
XX  
AC ADH13940;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:3.  
XX  
ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
KW cervical carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN EP1170372-Al.  
XX  
PD 09-JAN-2002.  
XX  
PF 06-JUL-2000; 2000EP-00114560.  
XX  
PR 06-JUL-2000; 2000EP-00114560.  
XX  
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
PI Roel F, Soto U, Coy J, Finzer P, Delius H, Poustka A;  
PI Zur Hausen H, Patzelt A;  
XX  
WPI; 2002-165895/22.  
PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
PT useful in gene therapy, for treating atherosclerosis and cancer.  
XX  
PS Disclosure; SEQ ID NO 3; 30pp; English.  
XX  
The invention relates to a novel nucleic acid molecule (I) comprising a  
sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
having the biological activity of MCP-1. A protein encoded by a nucleic  
acid of the invention has cytostatic, and antiarteriosclerotic activity.  
CC A nucleic acid of the invention may have a use in gene therapy. A  
compound of the invention is useful in the preparation of a medicament  
for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

Tue Aug 9 17:01:25 2005

CC identify a molecule or compound that specifically binds a cDNA of the  
CC invention. A protein encoded by the cDNA may be used to screen several  
CC molecules or compounds to identify a ligand that specifically binds to  
CC the protein, or to produce or purify an antibody to the protein that can  
CC be used to detect a protein in a sample or purify a natural or  
CC recombinant protein from a sample. The nucleotides may be useful for  
CC diagnosing, staging, treating, or monitoring the progression of treatment  
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery  
CC disease, hypertension, or stroke. The cDNAs can also be used for large-scale  
CC genetic or gene expression analysis of several new nucleic acid  
CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for  
CC diagnosing pre-pathologic disorders, and chronic or acute diseases  
CC associated with abnormalities in the expression, amount or distribution  
CC of the protein. The present sequence represents a cDNA of the invention  
CC that is differentially expressed in activated vascular tissue. Note: The  
CC sequence data for this patent did not form part of the specification, but  
CC was obtained in electronic format directly from USPTO at  
CC http.seqdata.uspto.gov/sequence.html?DocID=20020137081  
XX Sequence 150 BP; 34 A; 37 C; 23 G; 56 T; 0 U; 0 Other;  
SQ Query Match 100.0%; Score 150; DB 6; Length 150;  
Best Local Similarity 100.0%; Pred. No. 2.9e-34; Indels 0; Gaps 0;  
Matches 150; Conservative 0; Mismatches 0;  
QY 1 CAAAGATCACAATCTAGCTCTGAGGTATAGGAGGAGCACTGGGATTTAATGAGCTCTTT 60  
DB 1 CAAAGATCACAATCTAGCTCTGAGGTATAGGAGGAGCACTGGGATTTAATGAGCTCTTT 60  
QY 61 CTCTTCTCTGCTGCTGCTTTGCTTTTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 120  
DB 61 CTCTTCTCTGCTGCTGCTTTGCTTTTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 120  
QY 121 ATATCCAGTTCATCTCAAAATGCTTTTTC 150  
DB 121 ATATCCAGTTCATCTCAAAATGCTTTTTC 150  
RESULT 2  
ID ABX63839 standard; cDNA; 724 BP.  
XX AC ABX63839;  
XX DT 26-FEB-2003 (first entry)  
XX DE Human cDNA #839 differentially expressed in activated vascular tissue.  
XX KW Human; gene; ss; vascular tissue; cytostatic; atherosclerosis; cardiant;  
XX KW hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;  
XX KW gene therapy; vascular disease; cancer; coronary; artery disease;  
XX KW hypertension; diabetes; pre-eclampsia; restenosis;  
XX KW ischaemia-reperfusion injury; stroke.  
XX OS Homo sapiens.  
XX PN US2002137081-A1.  
XX PD 26-SEP-2002.  
XX PF 08-JAN-2002; 2002US-00044090.  
XX PR 28-JUL-2000; 2000US-0222469P.  
XX PR 08-JAN-2001; 2001US-0260483P.  
XX PA (BAND/) BANDMAN O.  
XX PI Bandman O;  
XX DR WPI; 2003-110597/10.  
XX PT Combination for diagnosing, staging, treating, or monitoring the  
XX PT progression of treatment of a vascular disease, e.g. atherosclerosis,  
XX PT comprises several cDNAs that are differentially expressed in activated  
XX PT vascular tissue.  
XX PS Claim 1; Page; 18pp; English.  
XX CC This invention relates to a combination comprising several cDNAs that are  
XX CC differentially expressed in activated vascular tissue. The invention also  
XX CC discloses a high throughput method for detecting differentially expressed  
XX CC cDNAs in a sample. The cDNAs of the invention may have  
XX CC antiatherosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;  
XX CC gynaecological; vasotropic and cerebroprotective activities and may be  
XX CC used in gene therapy. The cDNAs of the invention may be used in a high-  
XX CC throughput methods for detecting differential expression of one or more  
XX CC cDNAs in a sample, or screening several molecules or compounds to

CC identify a molecule or compound that specifically binds a cDNA of the  
CC invention. A protein encoded by the cDNA may be used to screen several  
CC molecules or compounds to identify a ligand that specifically binds to  
CC the protein, or to produce or purify an antibody to the protein that can  
CC be used to detect a protein in a sample or purify a natural or  
CC recombinant protein from a sample. The nucleotides may be useful for  
CC diagnosing, staging, treating, or monitoring the progression of treatment  
CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery  
CC disease, hypertension, or stroke. The cDNAs can also be used for large-scale  
CC genetic or gene expression analysis of several new nucleic acid  
CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for  
CC diagnosing pre-pathologic disorders, and chronic or acute diseases  
CC associated with abnormalities in the expression, amount or distribution  
CC of the protein. The present sequence represents a cDNA of the invention  
CC that is differentially expressed in activated vascular tissue. Note: The  
CC sequence data for this patent did not form part of the specification, but  
CC was obtained in electronic format directly from USPTO at  
CC http.seqdata.uspto.gov/sequence.html?DocID=20020137081  
XX Sequence 724 BP; 184 A; 180 C; 154 G; 204 T; 0 U; 2 Other;  
SQ Query Match 100.0%; Score 150; DB 8; Length 724;  
Best Local Similarity 100.0%; Pred. No. 4.4e-34; Indels 0; Gaps 0;  
Matches 150; Conservative 0; Mismatches 0;  
QY 1 CAAAGATCACAATCTAGCTCTGAGGTATAGGAGGAGCACTGGGATTTAATGAGCTCTTT 60  
DB 308 CAAAGATCACAATCTAGCTCTGAGGTATAGGAGGAGCACTGGGATTTAATGAGCTCTTT 367  
QY 61 CTCTTCTCTGCTGCTGCTTTGCTTTTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 120  
DB 368 CTCTTCTCTGCTGCTGCTTTGCTTTTCTCATGACTCTTTTCTGCTCTTAAGATCAGA 427  
QY 121 ATATCCAGTTCATCTCAAAATGCTTTTTC 150  
DB 428 ATATCCAGTTCATCTCAAAATGCTTTTTC 457  
RESULT 3  
ID ADO03803 standard; DNA; 11793 BP.  
XX AC ADO03803;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human Cc12 gene and enhancer region DNA SeqID 4.  
XX KW human; ds; animal model; age-related macular degeneration; AMD;  
XX KW gene knockout; Cc12-deficient; Ccr2-deficient; drusen;  
XX KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;  
XX KW chorioidal neovascularisation; ophthalmological; gene therapy.  
XX OS Homo sapiens.  
XX PN WO2004041160-A2.  
XX PD 21-MAY-2004.  
XX PF 16-OCT-2003; 2003WO-US032933.  
XX PR 30-OCT-2002; 2002US-0422096P.  
XX PA (KENT) UNIV KENTUCKY RES FOUND.  
XX PI Ambati J;  
XX DR WPI; 2004-400512/37.  
XX PT Testing candidate drug for treating age-related macular degeneration, by  
XX PT administering drug to Cc12-deficient, Ccr2-deficient knockout mouse, and  
XX PT analyzing development or regression of drusen and/or lipofuscin

PT accumulation in eye.  
XX Disclosure; SEQ ID NO 4; 64pp; English.  
XX  
CC This invention relates to a novel methods and animal models for testing  
CC candidate drugs that can be used for the treatment or prevention of age-  
CC related macular degeneration (AMD). Specifically, it refers to  
CC administering a candidate drug to gene knockout mice, in particular Ccl2-  
CC deficient, Ccr2-deficient and/ or a Ccl2-deficient/Ccr2-deficient dual  
CC knockout mouse. The present invention describes analysing the knockout  
CC mouse eye for development or regression of drusen and/ or lipofuscin  
CC accumulation, as well as for the effect of a candidate drug on Bruch's  
CC membrane, retinal degeneration and/ or choroidal neovascularisation.  
CC Accordingly, such compositions exhibit ophthalmological activities and  
CC can be used for gene therapy purposes. This polynucleotide sequence is  
CC the human Ccl2 gene and enhancer region DNA of the invention.  
XX  
SQ Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;  
Query Match 100.0%; Score 150; DB 12; Length 11793;  
Best Local Similarity 100.0%; Pred. No. 8.8e-34; Indels 0; Gaps 0;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGATCACATTCTAGCTCTGAGGTATAGGCAGAGCACTGGGATTTAATGAGCTCTTT 60  
DB 8299 CAAAGATCACATTCTAGCTCTGAGGTATAGGCAGAGCACTGGGATTTAATGAGCTCTTT 8358  
QY 61 CTCCTCTCCTGCTGCTTTGCTTTTCTCTCATGACTCTTTCTGCTCTTAAGATCAGA 120  
DB 8359 CTCCTCTCCTGCTGCTTTGCTTTTCTCTCATGACTCTTTCTGCTCTTAAGATCAGA 8418  
QY 121 ATAATCCAGTTCATCTCAAAATGCTTTTC 150  
DB 8419 ATAATCCAGTTCATCTCAAAATGCTTTTC 8448  
RESULT 4  
ID ACF64399  
XX ACF64399 standard; DNA; 750 BP.  
XX  
AC ACF64399;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human MCP1 nucleotide sequence >MCP1\_02.  
XX  
KW Human; detection; computer-readable storage medium; polymorphic site;  
KW signal carrying data; data processing system; multiple sclerosis; gene;  
KW ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2003014319-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 07-AUG-2002; 2002WO-US025268.  
XX  
PR 07-AUG-2001; 2001US-0310741P.  
PR 24-SEP-2001; 2001US-0324790P.  
XX  
PA (DNAS-) DNA SCI INC.  
XX  
PI Jones HB, Xu H, White R, Rienhoff HY, Jin W, Natsoulis G;  
XX  
DR WPI; 2003-268196/26.  
XX  
PT New polynucleotide, useful for detecting loci associated with multiple  
PT sclerosis.  
XX  
PS Disclosure; Page 74-75; 93pp; English.  
XX  
CC The present invention describes an isolated polynucleotide (PN)

CC comprising: (a) a sequence comprising at least 15 contiguous nucleotides  
CC of a sequence comprising variant sequences (A) from Table 4 given in the  
CC specification; or (b) a sequence that is complementary to (A). Also  
CC described: (1) an array of (PN)s comprising two or more of the isolated  
CC (PN)s; (2) detecting a (PN) in an individual; (3) a computer-readable  
CC storage medium, where each record has a field identifying a base  
CC occupying a (PN) site and a location of the polymorphic site; and (4) a  
CC signal carrying data for access by an application program having executed  
CC on a data processing system. The (PN) can be used for detecting loci  
CC associated with multiple sclerosis. ACF64025 to ACF64424 represent  
CC sequences used in the exemplification of the present invention  
XX  
SQ Sequence 750 BP; 186 A; 199 C; 144 G; 219 T; 0 U; 2 Other;  
Query Match 99.7%; Score 149.6; DB 8; Length 750;  
Best Local Similarity 99.3%; Pred. No. 5.8e-34;  
Matches 149; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGATCACATTCTAGCTCTGAGGTATAGGCAGAGCACTGGGATTTAATGAGCTCTTT 60  
DB 157 CAAAGATCACATTCTAGCTCTGAGGTATAGGCAGAGCACTGGGATTTAATGAGCTCTTT 216  
QY 61 CTCCTCTCCTGCTGCTTTGCTTTTCTCTCATGACTCTTTCTGCTCTTAAGATCAGA 120  
DB 217 CTCCTCTCCTGCTGCTTTGCTTTTCTCTCATGACTCTTTCTGCTCTTAAGATCAGA 276  
QY 121 ATAATCCAGTTCATCTCAAAATGCTTTTC 150  
DB 277 ATAATCCAGTTCATCTCAAAATGCTTTTC 306  
RESULT 5  
ID ABK47979  
XX ABK47979 standard; DNA; 2775 BP.  
XX  
AC ABK47979;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Human small inducible cytokine A2 (SCYA2) genomic DNA.  
XX  
KW Human; small inducible cytokine A2; SCYA2; gene; ds; haplotype pair;  
KW haplotyping; atherosclerosis; antiarteriosclerotic; gene therapy;  
KW single nucleotide polymorphism; genotyping; drug screening;  
KW chromosome 17q11.2-q21.1.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 598..2080  
FT /tag= a  
FT /product= "Human SCYA2"  
XX  
PN WO200218413-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 28-AUG-2001; 2001WO-US026899.  
XX  
PR 28-AUG-2000; 2000US-0228496P.  
XX  
PA (GENA-) GENAISANCE PHARM INC.  
XX  
PI Anastasio AE, Finkel K, Koshy B, Kumar AM, Lee HH;  
XX  
DR WPI; 2002-339655/37.  
DR P-PSDB; AAU77179.  
XX  
PT New genetic variants having polymorphisms in the small inducible cytokine  
PT A1 (SCYA2) gene, useful for studying the function of SCYA2, and for  
PT treating disorders affected by expression or function of the SCYA2  
PT isogene.  
XX





of expression of the gene is indicative of inflammation in a tissue, an allergic (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX Sequence 227968 BP; 53185 A; 60440 C; 61985 G; 52358 T; 0 U; 0 Other;  
SQ

Query Match 23.3%; Score 35; DB 6; Length 227968;  
Best Local Similarity 57.9%; Pred. No. 10;  
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 44 GATTTAATGAGCTCTTCTCTCTCTCTGCTGCTTTTGGCTTTTTCCTCATGACTCTTTT 103  
DB 44522 GACTTCTGTAATGATCTCCAAACCTGTTCTTTTGGTTCTCTCTCTTTTCTCT 44463  
QY 104 CTGCTCTTAAGATCAGAAATAATCCAGTTTCATCTAAATGCTTTTC 150  
DB 44462 CTTCTCTCTAGTTATTTCCTCGATTTCAATCTACGATGCTTCTAC 44416

RESULT 11  
ADQ18538/c  
ID ADQ18538 standard; DNA; 227968 BP.  
XX  
AC ADQ18538;  
XX  
DT 26-AUG-2004 (first entry)  
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 1357.  
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
KW ds.  
XX Homo sapiens.  
XX WO2004048938-A2.  
XX  
XX 10-JUN-2004.  
XX  
XX 26-NOV-2003; 2003WO-US038193.  
XX  
XX 26-NOV-2002; 2002US-0429739P.  
XX  
XX (PROT-) PROTEIN DESIGN LABS INC.  
PA  
XX Aziz N, Ginsburg WM, Zlotnik A;  
XX  
XX WPI; 2004-441208/41.  
XX  
XX Early detection of soft tissue sarcoma comprises determining expression  
PT Of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.  
XX  
XX Example 2; SEQ ID NO 1357; 210pp; English.  
XX



Tue Aug 9 17:01:25 2005

us-09-899-276c-3.rng

XX WO200166735-A1.  
XX 13-SEP-2001.  
XX 07-MAR-2001; 2001WO-JP001752.  
XX 08-MAR-2000; 2000JP-00063959.  
XX (WELF-) WELFIDE CORP.  
XX Uno S, Asakura E, Naito K;  
XX WPI; 2001-602631/68.  
XX Polypeptides that control MTRF and MTRF gene in mast cells for use in a  
XX related medical area.  
XX Claim 5; Page 57-59; 65pp; Japanese.  
XX The invention relates to polypeptides, comprising at least an optionally  
XX mutated amino acid sequence AAG78184-AAG78187 and having c-kit gene  
XX promoter controlling activity, especially in mast cells. The polypeptides  
XX are useful in diagnosing disease associated with mast cells and  
XX controlling MTRF (microphthalmia-associated transcription factor) and the  
XX MTRF gene, especially in mast cells. The present sequence is that of a  
XX MTRF protein isoform related DNA sequence  
SQ Sequence 2125 BP; 607 A; 535 C; 520 G; 463 T; 0 U; 0 Other;  
Query Match 22.7%; Score 34; DB 4; Length 2125;  
Best Local Similarity 54.9%; Pred. No. 6.2;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
QY 27 ATAGGCGAGCACTGGGATTTAATGAGCTCTTTCTCTCTGCTGCTGCTTTGCTTT 86  
DB 2115 AAGCGAGAAGTCTGGATGCTTTCCGACCACTTTCTCTCTCTCTCTCTCTCTCT 2056  
QY 87 TTCCTCATGACTCTTTCTGCTCTTAAGATCAGAAATATCCAGTTATCTCTAAATGCTT 146  
DB 2055 TTCCT 1996  
QY 147 TT 148  
DB 1995 TT 1994

RESULT 15  
AAI64627/c  
ID AAI64627 standard; cDNA; 2196 BP.  
XX AAI64627;  
XX 19-DEC-2001 (first entry)  
XX Mouse MTRF-origin protein isoform encoding cDNA SEQ ID NO 5.  
XX Mouse; MTRF; origin; c-kit; mast cell;  
XX microphthalmia-associated transcription factor; ss.  
XX Mus musculus.  
XX Key Location/Qualifiers  
XX CDS 59..1663  
XX /\*tag= a  
XX /product= "MTRF"  
XX WO200166735-A1.  
XX 13-SEP-2001.  
XX 07-MAR-2001; 2001WO-JP001752.  
XX

PR 08-MAR-2000; 2000JP-00063959.  
XX (WELF-) WELFIDE CORP.  
XX Uno S, Asakura E, Naito K;  
XX WPI; 2001-602631/68.  
XX P-ESDB; AAG78184.  
XX Polypeptides that control MTRF and MTRF gene in mast cells for use in a  
XX related medical area.  
XX Claim 3; Page 43-47; 65pp; Japanese.  
XX The invention relates to polypeptides, comprising at least an optionally  
XX mutated amino acid sequence AAG78184-AAG78187 and having c-kit gene  
XX promoter controlling activity, especially in mast cells. The polypeptides  
XX are useful in diagnosing disease associated with mast cells and  
XX controlling MTRF (microphthalmia-associated transcription factor) and the  
XX MTRF gene, especially in mast cells. The present sequence is that of a  
XX MTRF protein isoform encoding cDNA  
SQ Sequence 2196 BP; 627 A; 555 C; 535 G; 479 T; 0 U; 0 Other;  
Query Match 22.7%; Score 34; DB 4; Length 2196;  
Best Local Similarity 54.9%; Pred. No. 6.3;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
QY 27 ATAGGCGAGCACTGGGATTTAATGAGCTCTTTCTCTCTCTCTCTCTCTCTCTCTCT 86  
DB 2186 AAGCGAGAAGTCTGGATGCTTTCCGACCACTTTCTCTCTCTCTCTCTCTCTCTCT 2127  
QY 87 TTCCTCATGACTCTTTCTGCTCTTAAGATCAGAAATATCCAGTTATCTCTAAATGCTT 146  
DB 2126 TTCCT 2067  
QY 147 TT 148  
DB 2066 TT 2065  
Search completed: August 4, 2005, 08:31:10  
Job time : 104.413 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 707.976 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-3  
Perfect score: 150  
Sequence: 1 caaagatcacatttagctc.....tcactctaaatgcttttc 150

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37	24.7	750	4	BG109427 602280491
C 2	36.8	24.5	750	5	BH355897 CH230-139
C 3	36.2	24.1	660	8	AZ117778 RPCI-23-4
C 4	36	24.0	522	8	AQ994475 RPCI-23-2
C 5	36	24.0	564	8	CE774603 tigr-gss-
C 6	36	24.0	706	9	CV262112 WS02018 B
C 7	36	24.0	706	9	CV262112 WS02018 B
C 8	35.8	23.9	643	7	CC927097 T093121ba
C 9	35.6	23.7	740	9	BH355897 CH230-139
C 10	35.6	23.7	749	5	BH355897 CH230-139
C 11	35.4	23.6	434	7	CK094148 1014P80.3
C 12	35.4	23.6	431	7	CV040698 4138247 B
C 13	35.4	23.6	546	6	CD731079 4039997 1
C 14	35.4	23.6	626	5	BU492000 604131767
C 15	35.4	23.6	638	5	BX269592 BX269592
C 16	35.4	23.6	641	4	BM486827 pgm2n.pk0
C 17	35.4	23.6	641	5	BU431927 603831037
C 18	35.4	23.6	644	5	BU262950 603816306
C 19	35.4	23.6	655	7	CO140850 EST835521
C 20	35.4	23.6	662	5	BU276566 603866716
C 21	35.4	23.6	684	7	CO141545 EST836216
C 22	35.4	23.6	695	7	CO138311 EST832982
C 23	35.4	23.6	703	5	BU401493 604140230
C 24	35.4	23.6	706	5	BX277162 BX277162

C 25	35.4	23.6	726	5	BU293182 603607650
C 26	35.4	23.6	731	5	BU353399 603526962
C 27	35.4	23.6	733	5	BU445628 603766792
C 28	35.4	23.6	749	5	BU382780 603857431
C 29	35.4	23.6	753	5	BU368195 603567811
C 30	35.4	23.6	755	5	BU307556 603537684
C 31	35.4	23.6	758	5	BU280652 603862586
C 32	35.4	23.6	760	5	BU318233 603853088
C 33	35.4	23.6	779	5	BU248832 603778214
C 34	35.4	23.6	780	5	BU471485 60364739
C 35	35.4	23.6	786	5	BU121951 603146909
C 36	35.4	23.6	790	5	BU319724 603486583
C 37	35.4	23.6	792	5	BU420636 603957371
C 38	35.4	23.6	800	5	BU308950 603538608
C 39	35.4	23.6	807	5	BU372054 603810786
C 40	35.4	23.6	820	5	BU309968 603537452
C 41	35.4	23.6	845	5	BU123687 603147705
C 42	35.4	23.6	847	5	BU490127 604131257
C 43	35.4	23.6	863	5	BU330393 603496971
C 44	35.4	23.6	876	5	BU279119 603598901
C 45	35.4	23.6	907	5	BU387467 603857629

ALIGNMENTS

RESULT 1  
BG109427/c  
LOCUS 602280491F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4367999 5',  
DEFINITION mRNA sequence.  
ACCESSION BG109427  
VERSION BG109427.1 GI:12602933  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 750)  
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0021 row: e column: 24  
High quality sequence stop: 671.  
Location/Qualifiers  
1..750  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4367999"  
/tissue\_type="osteosarcoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 86"  
/note="Organ: bone; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally, oligo-dT primed. Average insert size 1.533 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

FEATURES  
source

ORIGIN

Query Match 24.7%; Score 37; DB 4; Length 750;  
Best Local Similarity 67.5%; Pred. No. 15;  
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

us-09-899-276c-3.rst

Tue Aug 9 17:01:26 2005

QY 40 CTGGATTAAATAGACTCTTTCTCTCTCTCCGCGCTGCTTTGCTTTTCTCTCATGACTC 99  
 |||||  
 Db 737 CTGACTATTAATCAGCACTTTGCCCTCTCTCCAGCAATGTATTATTTCTCCGATTC 678  
 |||||

QY 100 TTTTCTGCTCTTAAGAT 116  
 |||||  
 Db 677 TTTTCTGCTCTCTGT 661  
 |||||

RESULT 2  
 BU369171/c 750 bp mRNA linear EST 28-NOV-2002  
 LOCUS 603568157F1 CSEQCHN73 Gallus gallus cDNA clone ChEST524a16 5', mRNA  
 DEFINITION sequence.  
 ACCESSION BU369171  
 VERSION BU369171.1 GI:25877172  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 750)  
 AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.  
 TITLE A Comprehensive Collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 2235534  
 PUBMED 12445392  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

source  
 1..750  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Compton Line 151"  
 /db\_xref="taxon:9031"  
 /clone="ChEST524a16"  
 /sex="Female"  
 /tissue\_type="not cerebrum or cerebellum"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQCHN73"  
 /note="Organ: brain; Vector: pBluescript II KS(+); Site 1:  
 EcoRI; Site 2: NotI; This normalized library was  
 constructed from 1 million independent clones. cDNA  
 synthesis was initiated using an oligo(dT) primer, using  
 methylated C in the first strand synthesis reaction.  
 Following this first strand reaction, double-stranded cDNA  
 was blunted, ligated to NotI adapters, digested with  
 EcoRI, size-selected, and cloned into the NotI and EcoRI  
 compatible sites of a custom modified MCS of the  
 pBluescript (KS+) vector. The library was normalized in 2  
 rounds using conditions adapted from Soares et al., PNAS  
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
 (1996): 791, except that a significantly longer  
 reannealing hybridization was used."

## ORIGIN

Query Match 24.5%; Score 36.8; DB 5; Length 750;  
 Best Local Similarity 56.7%; Pred. No. 17;  
 Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
 QY 31 GCAGACACTGGGATTAATAGACTCTTTCTCTCTCTGCTGCTGCTTTGCTTTTCC 90  
 |||||  
 Db 392 GCAGAACTCAGGCTCTTAACAGGACAGTCTCTGTGCTGCTGATGTCAAGTATGAGGTG 333  
 |||||

QY 91 TCATGACTCTTTTCTGCTCTTAAGATCAGATAATCCAGTTCATCTCAATAATGCTTTTC 150  
 |||||  
 Db 332 TGATCAGTCGCTTCAACTCCTTGAAATAAAGCCCATCTCTTCTTAAGTTTTTTTCC 273  
 |||||

## RESULT 3

BH355897 660 bp DNA linear GSS 03-DEC-2001  
 LOCUS CH230-199118.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 DEFINITION CH230-199118, genomic survey sequence.

## ACCESSION

BH355897

## VERSION

BH355897.1 GI:17286631

## KEYWORDS

Rattus norvegicus (Norway rat)

## SOURCE

Rattus norvegicus

## ORGANISM

Rattus.

## REFERENCE

1 (bases 1 to 660)

## AUTHORS

Zhao,S.,Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,

## TITLE

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

## JOURNAL

Unpublished (1999)

## COMMENT

Other GSSs: CH230-199118.TV

## JOURNAL

Department of Eukaryotic Genomics

## MEDLINE

The Institute for Genomic Research

## PUBMED

9712 Medical Center Dr., Rockville, MD 20850, USA

## COMMENT

Tel: 301 838 0200

## FEATURES

Fax: 301 838 0208

## source

Email: szhao@tigr.org

## Location/Qualifiers

Clones are derived from the rat BAC library CHORI-230

## FEATURES

Clones are derived from the rat BAC library CHORI-230

## source

(http://www.chori.org/bacpac/rat230.htm). For BAC library

## FEATURES

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

## source

Clones may be purchased from BACPAC Resources

## FEATURES

(http://www.chori.org/bacpac/or ering information.htm). BAC end

## source

page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html

## FEATURES

Plate: 199 row: 1 column: 18

## source

Seq primer: SP6

## FEATURES

Class: BAC ends.

## Location/Qualifiers

1..660

## FEATURES

/organism="Rattus norvegicus"

## source

/mol\_type="genomic DNA"

## FEATURES

/strain="BN/SSHsd/MCW"

## source

/db\_xref="taxon:10116"

## FEATURES

/clone="CH230-199118"

## source

/sex="Female"

## FEATURES

/cell\_type="Brain"

## source

/clone\_lib="CHORI-230 Segment 1"

## FEATURES

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by  
 Pieter de Jong"

## ORIGIN

Query Match 24.1%; Score 36.2; DB 8; Length 660;  
 Best Local Similarity 62.9%; Pred. No. 24;  
 Matches 56; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
 QY 55 CTCTTTCTCTTCCTCGCTGCTTTTGGCTTTTCTCTCATGACTCTTTTCTGCTCTTAAG 114  
 |||||  
 Db 347 CTCATTCCTCCCTACACACCTCTTTGCTCTAATCTCATGCGCCATTTCTAGCTTAAG 406  
 |||||

QY 115 ATCAGATATATCCAGTTCATCTCAATAATG 143  
 |||||  
 Db 407 AAACATGATATTCGAACTCAACTTAATG 435  
 |||||

## RESULT 4

AZ117778 522 bp DNA linear GSS 12-MAY-2000  
 LOCUS RPCI-23-478J14.TJ RPCI-23 Mus musculus genomic clone  
 DEFINITION

ACCESSION  
 AZ117778  
 VERSION  
 AZ117778.1 GI:7781248  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 522)  
 REFERENCE  
 AUTHORS  
 Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S.,  
 Akimret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de  
 Jong, P. and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-23  
 Unpublished (1999)  
 Other GSSs: RPCI-23-478J14.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)  
 or from Resea ch Genetics ([info@resgen.com](http://info@resgen.com)). BAC end page:  
[http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 478 row: J column: 14  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES source

Location/Qualifiers  
 1..522  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-478J14"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"  
 /notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site: 1:  
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

## ORIGIN

Query Match 24.0%; Score 36; DB 8; Length 522;  
 Best Local Similarity 53.6%; Pred. No. 26;  
 Matches 75; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
 QY 11 ATTCTAGCTCTGAGGTATAGGCAGACCTGGGATTTAATGAGCTCTTTCTCTCTCT 70  
 Db 337 AGTCATCTCTGAGGCATGTTATATTAAATACCTTTCTGCTCTCTTATGCTTGT 278  
 QY 71 GCCTGCTTTTGTCTTCTCTATGACTCTTTTCTGCTCTTAAGATCAGAATAATCCAGT 130  
 Db 277 GGAAGTTTTTGTCTTCTTTATGTTGGCCTTTATAAGCTAATAATTTAATACCAGCCAGT 218  
 QY 131 TCATCCTAAATGCTTTTTC 150  
 Db 217 TCATTACAATATTTTATC 198

RESULT 5  
 AQ994475/c  
 LOCUS  
 DEFINITION  
 AQ994475  
 RPCI-23-270M5.TJ RPCI-23 Mus musculus genomic clone RPCI-23-270M5,  
 genomic survey sequence.  
 accession  
 AQ994475

# VERSION KEYWORDS SOURCE ORGANISM

AQ994475.1 GI:7069572  
 GSS  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 564)  
 REFERENCE  
 AUTHORS  
 Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S.,  
 Akimret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de  
 Jong, P. and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-23  
 Unpublished (1999)  
 Other GSSs: RPCI-23-270M5.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)  
 or from Resea ch Genetics ([info@resgen.com](http://info@resgen.com)). BAC end page:  
[http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 270 row: M column: 5  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES source

Location/Qualifiers  
 1..564  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-270M5"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"  
 /notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site: 1:  
 EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBACe3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

## ORIGIN

Query Match 24.0%; Score 36; DB 8; Length 564;  
 Best Local Similarity 53.6%; Pred. No. 27;  
 Matches 75; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
 QY 11 ATTCTAGCTCTGAGGTATAGGCAGACCTGGGATTTAATGAGCTCTTTCTCTCTCT 70  
 Db 389 AGTCATCTCTGAGGCATGTTATATTAAATACCTTTCTGCTCTCTTATGCTTGT 330  
 QY 71 GCCTGCTTTTGTCTTCTCTATGACTCTTTTCTGCTCTTAAGATCAGAATAATCCAGT 130  
 Db 329 GGAAGTTTTTGTCTTCTTTATGTTGGCCTTTATAAGCTAATAATTTAATACCAGCCAGT 270  
 QY 131 TCATCCTAAATGCTTTTTC 150  
 Db 269 TCATTACAATATTTTATC 250

RESULT 6  
 CE406656  
 LOCUS  
 DEFINITION  
 CE406656  
 tigr-gss-dog-17000335613614 Dog Library Canis familiaris genomic,  
 genomic survey sequence.  
 accession  
 CE406656  
 version  
 CE406656.1 GI:36659452  
 keywords  
 GSS.

us-09-899-276c-3.rst

Tue Aug 9 17:01:26 2005

SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE 1 (bases 1 to 706)  
 AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
 Venter, J.C.  
 TITLE The dog genome: survey sequencing and comparative analysis  
 JOURNAL Science 301 (5641), 1898-1903 (2003)  
 MEDLINE 22875432  
 PUBMED 14512627  
 COMMENT Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirknes@tigr.org  
 Class: shotgun.  
 FEATURES Location/Qualifiers  
 source 1..706  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"  
 /note="Site 1: BatXi; Libraries were prepared from  
 peripheral blood"  
 ORIGIN  
 Query Match 24.0%; Score 36; DB 9; Length 706;  
 Best Local Similarity 55.6%; Pred. No. 27;  
 Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
 QY 4 AGATCACATTCTAGCTCTGAGGTATAGGCAGACATGGGATTAATGAGCTCTTCTC 63  
 Db 81 AATGGTATCATAGCACATATAGTGTGTATACACCTTCTCTCTTTTGTCTTCTT 140  
 QY 64 TTCTCTGCTGCTGCTTTTCTTTTCTCATGACTCTTTCTGCTCTTAAGATCAGATA 123  
 Db 141 TTTTATCCCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 200  
 QY 124 ATCC 127  
 Db 201 TTCC 204  
 RESULT 7  
 CE774603 706 bp DNA linear GSS 30-SEP-2003  
 LOCUS tigr-gss-dog-17000330621810 Dog Library Canis familiaris genomic,  
 DEFINITION genomic survey sequence.  
 ACCESSION CE774603.1 GI:37115367  
 VERSION CE774603  
 KEYWORDS GSS.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE 1 (bases 1 to 706)  
 AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
 Venter, J.C.  
 TITLE The dog genome: survey sequencing and comparative analysis  
 JOURNAL Science 301 (5641), 1898-1903 (2003)  
 MEDLINE 22875432  
 PUBMED 14512627  
 COMMENT Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200

Fax: 301-838-0208  
 Email: ekirknes@tigr.org  
 Class: shotgun.  
 FEATURES Location/Qualifiers  
 source 1..706  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"  
 /note="Site 1: BatXi; Libraries were prepared from  
 peripheral blood"  
 ORIGIN  
 Query Match 24.0%; Score 36; DB 9; Length 706;  
 Best Local Similarity 54.5%; Pred. No. 27;  
 Matches 72; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
 QY 6 ATCACATTCTAGCTCTGAGGTATAGGCAGACATGGGATTAATGAGCTCTTCTTCT 65  
 Db 428 ATCGCTTATCAGGTGGCAAGGGAAGAGATGATGAGATTTCCTCTTCTTCTTTC 487  
 QY 66 CTCTGCTGCTGCTTTTCTTCTCATGACTCTTTCTGCTCTTAAGATCAGATAAT 125  
 Db 488 TTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 547  
 QY 126 CCAGTTTCATCCT 137  
 Db 548 CTAATTGTCCT 559  
 RESULT 8  
 CV262112/c 643 bp mRNA linear EST 22-SEP-2004  
 LOCUS WS02018.B21\_K16\_PT-XN-IB-N-A-11 Populus balsamifera subsp.  
 DEFINITION trichocarpa x Populus nigra cDNA clone WS02018\_K16 3', mRNA  
 sequence.  
 ACCESSION CV262112.1 GI:52515087  
 VERSION CV262112  
 KEYWORDS EST.  
 SOURCE Populus balsamifera subsp. trichocarpa x Populus nigra  
 ORGANISM Populus balsamifera subsp. trichocarpa x Populus nigra  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
 REFERENCE 1 (bases 1 to 643)  
 AUTHORS Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y.,  
 Kirkpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G.,  
 Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A.,  
 Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A.,  
 Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and  
 Bohlmann, J.  
 TITLE The poplar transcriptome: Analysis of expressed sequence tags from  
 multiple cDNA libraries  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Joerg Bohlmann  
 Genome BC forest genomics program  
 University of British Columbia  
 USC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,  
 Vancouver, British Columbia, Canada, V6T 1Z3  
 Tel: 1-604-822-0282  
 Fax: 1-604-822-6097  
 Email: bohlmann@interchange.ubc.ca  
 Plate: WS02018 row: K column: 16  
 High quality sequence stop: 643  
 POLYA=Yes.  
 FEATURES Location/Qualifiers  
 source 1..643  
 /organism="Populus balsamifera subsp. trichocarpa x  
 Populus nigra"  
 /mol\_type="mRNA"  
 /cultivar="NXM6"  
 /db\_xref="taxon:293756"



Tue Aug 9 17:01:26 2005

```

Db      268 CATTGAAGTGGAGAGGAGTACCTGAGAACTGACCTTTTCTGTTCTGAGCTCTGAGCTTC 327
QY      70 TCCCTGCCCTTTGCTTTTCTCAAGACTCTTTTCTGCTCTTAAGATCAGATAATCCAG 129
Db      328 CTCACTTCTTAACTGTTTCTTAGTTCCTCTCTTTTGACATATCTGGGGTTTITAGTA 387

QY      130 TTCTATCTTAATGCTTT 147
Db      388 TTCTCTCTTAATGTTGT 405

RESULT 11
LOCUS   CK094148      434 bp      mRNA      linear      EST 01-DEC-2003
DEFINITION
cDNA clone 1014P80 3', mRNA sequence.
ACCESSION
CK094148
VERSION
CK094148.1 GI:38578473
KEYWORDS
EST.
SOURCE
Populus tremula
ORGANISM
Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Malpighiales; Salicaceae; Populus.
1 (bases 1 to 434)
Sterky,F., Bhalerai,R., Unneberg,P., Segerman,B., Nilsson,P.,
Brunner,A.M., Campaa,L., Jonsson-Lindvall,J., Tandre,K.,
Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerai,R.P.,
Nilsson,O., Sandberg,G., Karlsson,J., Lundeborg,J. and Jansson,S.
A Populus EST resource for functional genomics
Unpublished (2003)
Other ESTs: 1014P80P
Contact: Bo Segerman
Umea Plant Science Center, Department of Plant Physiology
Umea University
901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.

FEATURES
source
1..434
/organism="Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:113636"
/clone="1014P80"
/tissue_type="Senescing leaves"
/clone_lib="Populus senescing leaves cDNA library"

ORIGIN
Query Match 23.6%; Score 35.4; DB 7; Length 434;
Best Local Similarity 61.3%; Pred. No. 38; Indels 0; Gaps 0;
Matches 57; Conservative 0; Mismatches 36;

QY      21 TGAGGTATAGCAGAACCTGGGATTAATAGCTCTTTCTCTCTCTGCTGCTGCTTT 80
Db      98 TGTGTACGAGAAATGCGGGAATAAAGAGCTCTTTCTTTCTGCTGCTTACTGTT 39

QY      81 TGCTTTTTCATGACTCTTTTCTGCTCTAA 113
Db      38 TAGTTTCTCCCATATCTCTGTCCAGTTTCA 6

RESULT 12
LOCUS   CV040698/c      491 bp      mRNA      linear      EST 01-OCT-2004
DEFINITION
3GAL 27P09 5', mRNA sequence.
ACCESSION
CV040698
VERSION
CV040698.1 GI:53560137
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 491)
Evock-Clover,C.M., Ashwell,C.M., McMurtry,J.P., Lillehoj,H.S.,
Matukumalli,L.K. and Van Tassel,C.P.
Characterization of expressed sequence tags generated from multiple
chicken tissues
Unpublished (2004)
Contact: Christina M. Clover
Growth Biology Laboratory
Animal and Natural Resources Institute
Bldg. 200 Rm215 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048224
Fax: 3015048223
Email: chrisceanri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt,, -trim fastavector identified by
cross match using options -minmatch 12 -minscore 12
Plate: 27 row: P column: 09
Seq primer: CCCAGTCACGAGCTGTATAACG
High quality sequence stop: 491.
Location/Qualifiers
1..491
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Leghorn and broiler"
/db_xref="taxon:9031"
/clone="3GAL 27P09"
/lab_host="DH5alpha"
/clone_lib="BARC 3GAL chicken mixed tissue"
/notes="Vector: pBluescript, SK+, Stratagene; Site_1: NotI;
Site_2: EcoRI; Normalized library of pooled RNA isolated
from whole brain, ultimobranchial gland, parathyroid
gland, cecal tonsil and primordial germ cells Multiple"

ORIGIN
Query Match 23.6%; Score 35.4; DB 7; Length 491;
Best Local Similarity 56.4%; Pred. No. 38; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 51;

QY      31 GCAGAGCACTGGGATTAATAGCTCTTTCTCTCTCTGCTGCTGCTTTTTC 90
Db      191 GCAGAAACTCAGGGCTCTAAACAGGACAGTCTTGTGCTGCTGCTGCTGCTG 132

QY      91 TCATGACTCTTTTCTGCTCTTAAGATCAGATAATCCAGTTCATCTTAAATGCTTT 147
Db      131 TGATCATGCTCGTTCACTTCTTGAAAAATAAAACGCCATTCCTTCTTAAGTTT 75

RESULT 13
LOCUS   CD731079/c      546 bp      mRNA      linear      EST 26-JUN-2003
DEFINITION
clone 1GAL_7D16 5', mRNA sequence.
ACCESSION
CD731079
VERSION
CD731079.1 GI:32281928
KEYWORDS
EST.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 546)
Min,W., Lillehoj,H.S., Ashwell,C.M., Matukumalli,L.K., van
Tassel,C. and Han,J.Y.
Chicken intestinal lymphocyte EST database as a resource for the
analysis of mucosal immune function
Unpublished (2003)
Contact: Hyun S. Lillehoj
Animal Parasite Diseases Laboratory
Animal and Natural Resources Institute, USDA
Bldg.1043, BARC-East, Beltsville, MD 20705, USA
Tel: 3015048771

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```



Tue Aug 9 17:01:26 2005

us-09-899-276c-3.rst

```
/dev stage="from embryos to adults"
/lab host="DH10B"
/clone_lib="AGENAE Gallus multi-tissues normalized
library (gcag)"
/note="Vector: pRT3D-pac; tissues: brain, embryos,
kidney, multi-tissues, muscle, pancreas, skin, testis,
liver, adipose tissue, granulosa, utero-vaginal gland,
oviduct, small follicle, ovary, hypothalamus, pituitary
gland, ileon, jejunum, caecum, duodenum, spleen,
fabricius gland, bone marrow, thymus, hematopoietic
progenitor cells. Clone distribution : AGENAE Resource
centre, Francois Piumi, Francois.Piumi.inra.fr, INRA, CEA
Radiobiologie et Etude du genome (LREG), Domaine de
Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"
```

ORIGIN

```
Query Match      23.6%; Score 35.4; DB 5; Length 638;
Best Local Similarity 56.4%; Pred. No. 40;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 31 GCAGAGACACTGGGATTTAATGAGCTCTTCTCTCTCTCTGCTGCTTTTCTTTTCC 90
Db 147 GCAGAACTCAGGGCTCTTAACAGGACAGCTCTCTGCTGCTGCTGCTCATGTCAGTTCATGAGGTG 88

QY 91 TCATGACTCTTTTCTGCTCTTAAGATCAGAAATATCCAGTTTCATCTTAAATGCTTT 147
Db 87 TGATCAGTCGGTTCAACTCCTTGAAAAATAAAACGCCATTCTCTTCTTAAAGTTTTTT 31
```

Search completed: August 4, 2005, 14:32:13  
Job time : 715.976 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 30.9109 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-3  
Perfect score: 150  
Sequence: 1 caaagatcacatttagctc.....tcattctaaagtcttttc 150

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	34.6	23.1	1664976	4 US-08-916-421B-1	Sequence 1, Appli
C 2	34.6	23.1	1664976	4 US-09-692-570-1	Sequence 1, Appli
C 3	33.8	22.5	112705	4 US-09-949-016-15830	Sequence 15630, A
4	33.4	22.3	601	4 US-09-949-016-152710	Sequence 152710, A
5	33.4	22.3	155019	4 US-09-949-016-16029	Sequence 16029, A
C 6	33.2	22.1	100463	4 US-09-949-016-12511	Sequence 12511, A
C 7	33.2	22.1	100468	4 US-09-949-016-13725	Sequence 13725, A
8	32.2	21.5	7218	1 US-08-232-463-14	Sequence 14, Appl
9	32.2	21.5	167708	4 US-09-949-016-16423	Sequence 16423, A
C 10	32	21.3	36755	4 US-09-949-016-16994	Sequence 16994, A
C 11	32	21.3	38059	4 US-09-328-925-4	Sequence 4, Appli
C 12	31.8	21.2	48691	4 US-09-949-016-16308	Sequence 16308, A
13	31.8	21.2	194537	4 US-09-949-016-12828	Sequence 12828, A
14	31.8	21.2	201529	4 US-09-949-016-12740	Sequence 12740, A
15	31.6	21.1	4745	4 US-09-949-016-4716	Sequence 4716, Ap
16	31.6	21.1	6232	4 US-09-949-016-191	Sequence 191, App
17	31.6	21.1	21451	4 US-09-949-016-16458	Sequence 16458, A
18	31.6	21.1	59240	4 US-09-949-016-11933	Sequence 11933, A
19	31.6	21.1	183202	4 US-09-949-016-13614	Sequence 13614, A
C 20	31.4	20.9	71119	4 US-09-949-016-15358	Sequence 15358, A
21	31.2	20.8	98567	4 US-09-949-016-11750	Sequence 11750, A
22	31.2	20.8	100567	4 US-09-949-016-16934	Sequence 16934, A
C 23	31	20.7	1225	1 US-08-181-629A-3	Sequence 3, Appli
C 24	31	20.7	5496	1 US-08-181-629A-2	Sequence 2, Appli
25	31	20.7	52992	4 US-09-949-016-16105	Sequence 16105, A
C 26	30.8	20.5	601	4 US-09-949-016-70039	Sequence 70039, A
27	30.8	20.5	70088	4 US-09-949-016-16845	Sequence 16845, A

C 28	30.8	20.5	168971	4 US-09-949-016-13807	Sequence 13807, A
29	30.8	20.5	194937	4 US-09-949-016-17032	Sequence 17032, A
C 30	30.8	20.5	194937	4 US-09-949-016-17033	Sequence 17033, A
31	30.6	20.4	54033	4 US-09-949-016-12091	Sequence 12091, A
32	30.6	20.4	54033	4 US-09-949-016-14325	Sequence 14325, A
C 33	30.6	20.4	60417	4 US-09-949-016-13312	Sequence 13312, A
C 34	30.4	20.3	601	4 US-09-949-016-31004	Sequence 31004, A
C 35	30.4	20.3	601	4 US-09-949-016-31005	Sequence 31005, A
C 36	30.4	20.3	601	4 US-09-949-016-40864	Sequence 40864, A
C 37	30.4	20.3	601	4 US-09-949-016-40865	Sequence 40865, A
C 38	30.4	20.3	48763	4 US-09-916-204-3	Sequence 3, Appli
C 39	30.4	20.3	48763	4 US-10-282-048-3	Sequence 3, Appli
C 40	30.4	20.3	174493	4 US-09-804-471A-3	Sequence 3, Appli
C 41	30.4	20.3	174493	4 US-10-238-709-3	Sequence 3, Appli
42	30.4	20.3	462589	4 US-09-949-016-12900	Sequence 12900, A
43	30.4	20.3	476044	4 US-09-949-016-12412	Sequence 12412, A
C 44	30.2	20.1	601	4 US-09-949-016-48236	Sequence 48236, A
45	30.2	20.1	601	4 US-09-949-016-190132	Sequence 190132, A

ALIGNMENTS

RESULT 1

US-08-916-421B-1/c  
; Sequence 1, Application US/08916421B  
; Patent No. 6503729  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii  
; Patent No. 6503729  
; FILE REFERENCE: PB275  
; CURRENT APPLICATION NUMBER: US/08/916,421B  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)

OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (103998).. <td>(103998)</td>	(103998)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (148948).. <td>(148948)</td>	(148948)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (163385).. <td>(163385)</td>	(163385)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (191989).. <td>(191989)</td>	(191989)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (191995).. <td>(191995)</td>	(191995)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (231980).. <td>(231980)</td>	(231980)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (234187).. <td>(234187)</td>	(234187)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (234220).. <td>(234220)</td>	(234220)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (234814).. <td>(234814)</td>	(234814)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (309398).. <td>(309398)</td>	(309398)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (309418).. <td>(309418)</td>	(309418)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (312837).. <td>(312837)</td>	(312837)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (312993).. <td>(312993)</td>	(312993)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (319226).. <td>(319226)</td>	(319226)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (559167).. <td>(559167)</td>	(559167)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (559241).. <td>(559241)</td>	(559241)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (600992).. <td>(600992)</td>	(600992)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (622708).. <td>(622708)</td>	(622708)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (657081).. <td>(657081)</td>	(657081)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (657203).. <td>(657203)</td>	(657203)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (674435).. <td>(674435)</td>	(674435)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (682442).. <td>(682442)</td>	(682442)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (713652).. <td>(713652)</td>	(713652)
OTHER INFORMATION: n equals a, t, c, or g	
NAME/KEY: misc feature	
LOCATION: (741684).. <td>(741684)</td>	(741684)
OTHER INFORMATION: n equals a, t, c, or g	

```

; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g

```

Tue Aug 9 17:01:25 2005

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15630
; LENGTH: 112705
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(112705)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-15630

Query Match      22.5%; Score 33.8; DB 4; Length 112705;
Best Local Similarity 60.2%; Pred. No. 2.7; Indels 0; Gaps 0;
Matches 56; Conservative

QY      20 CTGAGGTATAGCGAGCAAGCACTGGGATTTAATGAGCTCTTTCTCTCTCCGCGCTT 79
DB      85636 CTAGAGGAAAACAAGGATCTGTGTTAAATCAGCTCAGACTCTCTCTCTCTTCCATT 85695

QY      80 TTGCTTTTCTCATGACTCTTTTCTGCTCTTA 112
DB      85696 CTTTCATATTCCTATGCTCTCTCTCTGCTCATA 85728

RESULT 4
US-09-949-016-152710
; Sequence 152710, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152710
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-152710

Query Match      22.3%; Score 33.4; DB 4; Length 601;
Best Local Similarity 58.6%; Pred. No. 0.45; Indels 0; Gaps 0;
Matches 58; Conservative

QY      46 TTTAATGAGCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105
DB      370 TTCACTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 429

QY      106 GCTCTTAAGATCAGATAATCCAGTTTCATCTAAATGC 144
DB      430 TCTTTATGGAGCTAAATTTCTTACGTCGTGCTAAATAC 468

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Query Match      23.1%; Score 34.6; DB 4; Length 1664976;
Best Local Similarity 59.1%; Pred. No. 4.4; Indels 44; Gaps 0;
Matches 61; Conservative

QY      46 TTTAATGAGCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 105
DB      397144 TTTATTTTATCTTTATCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 397085

QY      106 GCTCTTAAGATCAGATAATCCAGTTTCATCTAAATGCCTTTTC 150
DB      397084 ATCATCATATCATATAATCAATCGCAATATCATATCTCTTC 397040

RESULT 3
US-09-949-016-15630
; Sequence 15630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

<pre>Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0; Qy 41 TGGGATTTAATAGACGTCTTTCCTCCTCGCTGCCTTGCTTTTCCCTCATGACTCT 100 Db 98488 TGGGTTTTATTTCCTTCCTTCCTCTTCTCTCGCTTGTTGTTTTCTTCTCTGAGTGT 98429 Qy 101 TTTCCTGCTCT 110 Db 98428 GTCGGTCTAT 98419  RESULT 7 US-09-949-016-13725/c ; Sequence 13725, Application US/09949016 ; Patent No. 6812339 ; GENERAL INFORMATION: ; APPLICANT: VENTER, J. Craig et al. ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF ; FILE REFERENCE: CL001307 ; CURRENT APPLICATION NUMBER: US/09/949,016 ; PRIOR FILING DATE: 2000-04-14 ; PRIOR APPLICATION NUMBER: 60/241,755 ; PRIOR FILING DATE: 2000-10-20 ; PRIOR APPLICATION NUMBER: 60/237,768 ; PRIOR FILING DATE: 2000-10-03 ; PRIOR APPLICATION NUMBER: 60/231,498 ; PRIOR FILING DATE: 2000-09-08 ; NUMBER OF SEQ ID NOS: 207012 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 13725 ; LENGTH: 100468 ; TYPE: DNA ; ORGANISM: Human ; FEATURE: ; NAME/KEY: misc_feature ; LOCATION: (1)...(100468) ; OTHER INFORMATION: n = A,T,C or G US-09-949-016-13725</pre>	<pre>Query Match 22.1%; Score 33.2; DB 4; Length 100468; Best Local Similarity 67.1%; Pred. No. 4; Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0; Qy 41 TGGGATTTAATAGACGTCTTTCCTCCTCGCTGCCTTGCTTTTCCCTCATGACTCT 100 Db 98488 TGGGTTTTATTTCCTTCCTTCCTCTCTCGCTTGTTGTTTTCTTCTCTGAGTGT 98429 Qy 101 TTTCCTGCTCT 110 Db 98428 GTCGGTCTAT 98419  RESULT 8 US-08-232-463-14 ; Sequence 14, Application US/08232463 ; Patent No. 5670367 ; GENERAL INFORMATION: ; APPLICANT: DORNER, F. ; APPLICANT: SCHEIFLINGER, F. ; APPLICANT: FALKNER, F. G. ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS ; NUMBER OF SEQUENCES: 52 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: Foley &amp; Lardner ; STREET: 1800 Diagonal Road, Suite 500 ; CITY: Alexandria ; STATE: VA ; COUNTRY: USA ; ZIP: 22313-0299 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible</pre>
<pre>Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0; Qy 46 TTTTAATGAGCTCTTCTCTCTCTCTGCTGCTTTTGGCTTTTTCCTCATGACTCTTTCT 105 Db 98206 TTCACCTAGTTTGTTCCT 98265 Qy 106 GCCTTTAAATGATCAATAATCCAGTTTCATCTCAAATGC 144 Db 98266 TCCTTTATGAGCTAATTTCTTAGCTGCTGAATAATC 98304  RESULT 6 US-09-949-016-12511/c ; Sequence 12511, Application US/09949016 ; Patent No. 6812339 ; GENERAL INFORMATION: ; APPLICANT: VENTER, J. Craig et al. ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF ; FILE REFERENCE: CL001307 ; CURRENT APPLICATION NUMBER: US/09/949,016 ; PRIOR FILING DATE: 2000-04-14 ; PRIOR APPLICATION NUMBER: 60/241,755 ; PRIOR FILING DATE: 2000-10-20 ; PRIOR APPLICATION NUMBER: 60/237,768 ; PRIOR FILING DATE: 2000-10-03 ; PRIOR APPLICATION NUMBER: 60/231,498 ; PRIOR FILING DATE: 2000-09-08 ; NUMBER OF SEQ ID NOS: 207012 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 12511 ; LENGTH: 100463 ; TYPE: DNA ; ORGANISM: Human ; FEATURE: ; NAME/KEY: misc_feature ; LOCATION: (1)...(100463) ; OTHER INFORMATION: n = A,T,C or G US-09-949-016-12511</pre>	<pre>Query Match 22.1%; Score 33.2; DB 4; Length 100463; Best Local Similarity 67.1%; Pred. No. 4; Matches 47; Conservative 0; Mismatches 23; Indels 0; Gaps 0; Qy 41 TGGGATTTAATAGACGTCTTTCCTCCTCGCTGCCTTGCTTTTCCCTCATGACTCT 100 Db 98488 TGGGTTTTATTTCCTTCCTTCCTCTTCTCTCGCTTGTTGTTTTCTTCTCTGAGTGT 98429 Qy 101 TTTCCTGCTCT 110 Db 98428 GTCGGTCTAT 98419  RESULT 7 US-09-949-016-13725/c ; Sequence 13725, Application US/09949016 ; Patent No. 6812339 ; GENERAL INFORMATION: ; APPLICANT: VENTER, J. Craig et al. ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF ; FILE REFERENCE: CL001307 ; CURRENT APPLICATION NUMBER: US/09/949,016 ; PRIOR FILING DATE: 2000-04-14 ; PRIOR APPLICATION NUMBER: 60/241,755 ; PRIOR FILING DATE: 2000-10-20 ; PRIOR APPLICATION NUMBER: 60/237,768 ; PRIOR FILING DATE: 2000-10-03 ; PRIOR APPLICATION NUMBER: 60/231,498 ; PRIOR FILING DATE: 2000-09-08 ; NUMBER OF SEQ ID NOS: 207012 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 13725 ; LENGTH: 100468 ; TYPE: DNA ; ORGANISM: Human ; FEATURE: ; NAME/KEY: misc_feature ; LOCATION: (1)...(100468) ; OTHER INFORMATION: n = A,T,C or G US-09-949-016-13725</pre>

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match          21.5%; Score 32.2; DB 1; Length 7218;
Best Local Similarity 13.1%; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 19; Conservative 74; Mismatches 52;

QY 6 ATCACTTCTAGCTCTGAGGTATAGGCGAAGACACTGGGATTAAATGAGCTCTTTCTTT 65
Db 1026 ATTAATCCGAGCTTGCTGAGTGCAGGAGGAGCTTGCATGATGATGATGATGATGAT 1085

QY 66 CTCCTGCTGCTTTTCTGCTTTTCTCTCATGACTTTTCTGCTCTTAAGATCAGATAAT 125
Db 1086 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1145

QY 126 CCAGTTCATCTCTAAATGCTTTTC 150
Db 1146 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1170

RESULT 9
US-09-949-016-16423
; Sequence 16423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16423
; LENGTH: 167708
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16423

Query Match          21.5%; Score 32.2; DB 4; Length 167708;
Best Local Similarity 52.6%; Pred. No. 10; Indels 167708;
Matches 19; Conservative 74; Mismatches 52;

QY 123 AATC 126
Db 5468 TCTC 5465

RESULT 11
US-09-328-925-4/c
; Sequence 4, Application US/09328925
; Patent No. 6610906
; GENERAL INFORMATION:
; APPLICANT: Kurachi, Kotoku
; APPLICANT: Kurachi, Sumiko
; TITLE OF INVENTION: Nucleotide Sequences for Gene Regulation and Methods of
; Use Thereof
; FILE REFERENCE: UM-03603
; CURRENT APPLICATION NUMBER: US/09/328,925
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 38059

```



Search completed: August 4, 2005, 14:40:44  
Job time : 39.9109 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 267.571 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-3

Perfect score: 150

Sequence: 1 caaagatcacatttagctc.....tcctctaaatgctttttc 150

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150	100.0	150	9	US-09-899-276-3
2	150	100.0	724	13	US-10-044-090-839
3	150	100.0	11793	19	US-10-685-705-4
4	146.4	97.6	2776	22	US-10-833-656-2
c 5	111.6	74.4	5926	15	US-10-311-455-1626
6	92	61.3	5926	15	US-10-311-455-1625
c 7	35	23.3	227968	20	US-10-723-860-1357

8	33.6	22.4	81684	19	US-10-322-281-673	Sequence 673, App
c 9	33.4	22.3	603	13	US-10-027-632-105897	Sequence 105897,
c 10	33.4	22.3	603	13	US-10-027-632-105898	Sequence 105898,
c 11	33.4	22.3	603	13	US-10-027-632-105899	Sequence 105899,
c 12	33.4	22.3	603	17	US-10-027-632-105897	Sequence 105897,
c 13	33.4	22.3	603	17	US-10-027-632-105898	Sequence 105898,
c 14	33.4	22.3	603	17	US-10-027-632-105899	Sequence 105899,
c 15	33.4	22.3	754	13	US-10-027-632-19185	Sequence 19185, A
c 16	33.4	22.3	754	13	US-10-027-632-19186	Sequence 19186, A
c 17	33.4	22.3	754	13	US-10-027-632-19187	Sequence 19187, A
c 18	33.4	22.3	754	17	US-10-027-632-19185	Sequence 19185, A
c 19	33.4	22.3	754	17	US-10-027-632-19186	Sequence 19186, A
c 20	33.4	22.3	754	17	US-10-027-632-19187	Sequence 19187, A
c 21	33.4	22.3	2631	11	US-09-973-278-926	Sequence 926, App
c 22	33.2	22.1	276	20	US-10-425-115-8798	Sequence 8798, Ap
c 23	32.8	21.9	467	10	US-09-918-995-13313	Sequence 13313, A
c 24	32.6	21.7	60604	18	US-10-300-283-11	Sequence 11, Appl
c 25	32.4	21.6	575	13	US-10-027-632-224845	Sequence 224845,
c 26	32.4	21.6	575	17	US-10-027-632-224845	Sequence 224845,
c 27	32.4	21.6	637	13	US-10-027-632-234763	Sequence 234763,
c 28	32.4	21.6	637	13	US-10-027-632-234764	Sequence 234764,
c 29	32.4	21.6	637	17	US-10-027-632-234763	Sequence 234763,
c 30	32.4	21.6	637	17	US-10-027-632-234764	Sequence 234764,
c 31	32	21.3	407	9	US-09-880-107-3095	Sequence 3095, Ap
c 32	32	21.3	600	22	US-10-972-079-37294	Sequence 37294, A
c 33	32	21.3	600	22	US-10-972-079-37295	Sequence 37295, A
c 34	32	21.3	611	13	US-10-027-632-105792	Sequence 105792,
c 35	32	21.3	611	13	US-10-027-632-105793	Sequence 105793,
c 36	32	21.3	611	13	US-10-027-632-105794	Sequence 105794,
c 37	32	21.3	611	17	US-10-027-632-105792	Sequence 105792,
c 38	32	21.3	611	17	US-10-027-632-105793	Sequence 105793,
c 39	32	21.3	611	17	US-10-027-632-105794	Sequence 105794,
c 40	32	21.3	736	13	US-10-027-632-18618	Sequence 18618, A
c 41	32	21.3	736	13	US-10-027-632-18619	Sequence 18619, A
c 42	32	21.3	736	13	US-10-027-632-18620	Sequence 18620, A
c 43	32	21.3	736	17	US-10-027-632-18618	Sequence 18618, A
c 44	32	21.3	736	17	US-10-027-632-18619	Sequence 18619, A
c 45	32	21.3	736	17	US-10-027-632-18620	Sequence 18620, A

ALIGNMENTS

RESULT 1  
US-09-899-276-3  
; Sequence 3, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Dellus, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899, 276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 150  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-3

Query Match 100.0%; Score 150; DB 9; Length 150;  
Best Local Similarity 100.0%; Pred. No. 3.4e-37;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue Aug 9 17:01:26 2005

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-685-705-4
Query Match 100.0%; Score 150; DB 19; Length 11793;
Best Local Similarity 100.0%; Pred. No. 1.7e-36; Indels 0; Gaps 0;
Matches 150; Conservative 0; Mismatches 0;

Qy 1 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 60
Db 1 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 60
Qy 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 120
Db 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 120
Qy 121 ATAATCCAGTTCATCCTAAATGCTTTTTC 150
Db 121 ATAATCCAGTTCATCCTAAATGCTTTTTC 150

RESULT 2
US-10-044-090-839
; Sequence 839, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: FA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 839
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1075405.1
; NAME/KEY: unsure
; LOCATION: 677, 710
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-839
Query Match 100.0%; Score 150; DB 13; Length 724;
Best Local Similarity 100.0%; Pred. No. 6.1e-37;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 60
Db 308 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 367
Qy 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 120
Db 368 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 427
Qy 121 ATAATCCAGTTCATCCTAAATGCTTTTTC 150
Db 428 ATAATCCAGTTCATCCTAAATGCTTTTTC 457

RESULT 3
US-10-685-705-4
; Sequence 4, Application US/10685705
; Publication No. US20040177387A1
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: JAVAKRISHNA, Ambati
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular
; FILE REFERENCE: 050229-0415
; CURRENT APPLICATION NUMBER: US/10/685,705
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/422,096
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 11793
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-685-705-4
Query Match 100.0%; Score 150; DB 19; Length 11793;
Best Local Similarity 100.0%; Pred. No. 1.7e-36; Indels 0; Gaps 0;
Matches 150; Conservative 0; Mismatches 0;

Qy 1 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 60
Db 8299 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 8358
Qy 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 120
Db 8359 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 8418
Qy 121 ATAATCCAGTTCATCCTAAATGCTTTTTC 150
Db 8419 ATAATCCAGTTCATCCTAAATGCTTTTTC 8448

RESULT 4
US-10-833-656-2
; Sequence 2, Application US/10833656
; Publication No. US20050148507A1
; GENERAL INFORMATION:
; APPLICANT: Wandl, Robert
; APPLICANT: Necina, Roman
; APPLICANT: Doods, Henri
; APPLICANT: Lenter, Martin
; APPLICANT: Seidler, Randolph
; TITLE OF INVENTION: Method for the production of an N-terminally modified
; FILE REFERENCE: Case 1/1492
; CURRENT APPLICATION NUMBER: US/10/833,656
; CURRENT FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2776
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2582)..(2582)
; OTHER INFORMATION: n = c, a, t or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2583)..(2583)
; OTHER INFORMATION: n = c, a, t or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2585)..(2585)
; OTHER INFORMATION: n = c, a, t or g
US-10-833-656-2
Query Match 97.6%; Score 146.4; DB 22; Length 2776;
Best Local Similarity 99.3%; Pred. No. 1.4e-35;
Matches 147; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 60
Db 1304 CAAAGATCACATCTAGCTCTGAGGTATAGGAGAGCACTGGGATTTAATGAGCTCTTT 1363
Qy 61 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 120
Db 1364 CTCCTCTCTGCTGCTGCTTTTCTTTTCCCTCATGACTCTTTTCTGCTCTTTAAGATCAGA 1423
Qy 121 ATAATCCAGTTCATCCTAAATGCTTTT 148
Db 1424 ATAATCCAGTTCATCCTAAATGCTTTT 1451
```

```

Query Match          61.3%; Score 92; DB 15; Length 5926;
Best Local Similarity .76.4%; Pred. No. 2.5e-18;
Matches 113; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      2   AAAGATCACATTCTAGCTCTGAGGTATAGCAGAACACTGGGATTTTAATGAGCTCTTTC 61
Db      5758  AAAGATTATAATTTAGTTTTGAGGTATAGTAGAAGTATGGGATTTTAATGAGTTTTTG 5817

QY      62   TCTTCTCTGCCTGCCCTTTTGCTTTTTTCCCATGACTCTTTTCTGCTCTTAAGATCAGAA 121
Db      5818  TTTTCTTTTGTGTTTGTGTTTTGTTTTTTTTTTATGATTTTTTTTTTGTGTTTAAAGATTAGAA 5877

QY      122  TAATCCAGTTCATCCTAAAATGCTTTTT 149
Db      5878  TAATTTAGTTTATTTAAAAATGTTTTT 5905

RESULT 7
US-10-723-860-1357/c
; Sequence 1357, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 1357
; LENGTH: 227968
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1357

Query Match          23.3%; Score 35; DB 20; Length 227968;
Best Local Similarity 57.9%; Pred. No. 8.7;
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY      44   GATTTAATGAGCTCTTTCTCTCTCGCTGCTTTTGTCTTTTCTCATGACTCTTTT 103
Db      44522  GACTTCTGGATGATCTCCAAACCCTGTTTCTTTTGTCTTTTCTTTTCTTTTCT 44463

QY      104   CTGCTCTTAAGATCAAGATAATCCAGTTCATCTAAAATGCTTTTTC 150
Db      44462  CTTCTCTCTAGTTTATTTCCCTCGATTTCAAATCTACGATGCTTCTAC 44416

RESULT 8
US-10-322-281-673
; Sequence 673, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 673
; LENGTH: 81684
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(81684)

```

```
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-673

Query Match          22.4%; Score 33.6; DB 19; Length 81684;
Best Local Similarity 59.4%; Pred. No. 16;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy      15 TAGCTGTGAGTATAGCAGAACACTGGGATTAAATAGACTCTTTCTCTTCCTCGCT 74
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      51805 TATACCTTGAGTAACGGTAATCTACATCTCATATAATGGTCTCTTCCCTTCTCTTCT 51864
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      75 GCCTTTTGCTTTTTTCCTCATGACTCTTTTCTGCTCT 110
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      51865 TTGCTTTCTTTTCCCTCTTTCTCTTCCTTCCTCTCTCT 51900
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-027-632-105897/c
; Sequence 105897, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105897
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-105897

Query Match          22.3%; Score 33.4; DB 13; Length 603;
Best Local Similarity 58.6%; Pred. No. 3.1;
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      46 TTTAATGAGCTCTTTCTCTCTCCCTGCCCTTTTGCCTTTTCTCATGACTCTTTCT 105
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      219 TTCACATAGATTGTGTTTCTCTCTCCCTTCCTCCCTTCCTCCCCTTCCCTTCTCTCTCTTT 160
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy      106 GCTCTTAAGATCAGAATAATCCAGTTTCATCTCTAAAATGC 144
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      159 TCITTTATGGAGCTAATTTCTCAGTCTGGGCTAAAAATAC 121
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-10-027-632-105898/c
; Sequence 105898, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

Db 219 TTCAACTAGTTTGGTTTCTCTCTCCCTTCCACTCCCTTTCCTCTGTCCTCTCTTT 160  
QY 106 GCTCTTAAGATCAGATAATCCAGTTCATCCCTAAATGC 144  
Db 159 TCTTTATGAGCTAATTTCCCTACGCTCGCTAAATAATAC 121

RESULT 12

US-10-027-632-105897/c  
; Sequence 105897, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 105897  
; LENGTH: 603  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-105897

Query Match 22.3%; Score 33.4; DB 17; Length 603;  
Best Local Similarity 58.6%; Pred. No. 3.1;  
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 46 TTTAATGAGCTCTTTCTCTCTCCCTGCGCTTTTGCTTTTTCCTCATGACTCTTTCT 105  
Db 219 TTCAACTAGTTTGGTTTCTCTCTCCCTTCCTCCCTTCCTCTCTCTCTCTCTTT 160  
QY 106 GCTCTTAAGATCAGATAATCCAGTTCATCCCTAAATGC 144  
Db 159 TCTTTATGAGCTAATTTCCCTACGCTCGCTAAATAATAC 121

RESULT 13

US-10-027-632-105898/c  
; Sequence 105898, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 105898  
; LENGTH: 603  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-105898

Query Match 22.3%; Score 33.4; DB 17; Length 603;  
Best Local Similarity 58.6%; Pred. No. 3.1;  
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 46 TTTAATGAGCTCTTTCTCTCTCCCTGCGCTTTTGCTTTTTCCTCATGACTCTTTCT 105  
Db 219 TTCAACTAGTTTGGTTTCTCTCTCCCTTCCTCCCTTCCTCTCTCTCTCTTT 160  
QY 106 GCTCTTAAGATCAGATAATCCAGTTCATCCCTAAATGC 144  
Db 159 TCTTTATGAGCTAATTTCCCTACGCTCGCTAAATAATAC 121

RESULT 14

US-10-027-632-105899/c  
; Sequence 105899, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 105899  
; LENGTH: 603  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-105899

Query Match 22.3%; Score 33.4; DB 17; Length 603;  
Best Local Similarity 58.6%; Pred. No. 3.1;  
Matches 58; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 46 TTTAATGAGCTCTTTCTCTCTCCCTGCGCTTTTGCTTTTTCCTCATGACTCTTTCT 105  
Db 219 TTCAACTAGTTTGGTTTCTCTCTCCCTTCCTCCCTTCCTCTCTCTCTCTTT 160  
QY 106 GCTCTTAAGATCAGATAATCCAGTTCATCCCTAAATGC 144  
Db 159 TCTTTATGAGCTAATTTCCCTACGCTCGCTAAATAATAC 121

Tue Aug 9 17:01:26 2005

RESULT 15  
US-10-027-632-19185/c  
; Sequence 19185, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 19185  
; LENGTH: 754  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-19185

Query Match 22.3%; Score 33.4; DB 13; Length 754;  
Best Local Similarity 58.6%; Pred. No. 3.4; Indels 0; Gaps 0;  
Matches 58; Conservative 0; Mismatches 41;  
QY 46 TTATATGAGCTTTCTCTTCCTGCTGCTTTTGTCTTTTCCTCATGACTCTTTCT 105  
Db 220 TTCAACTAGTTTGTCT 161  
QY 106 GCTCTTAAGATCAGATAATTCAGTTTCATCTAAATGC 144  
Db 160 TCUTTTATGGAGCTAATTTCTCTACGCTGGCTAAATAC 122

Search completed: August 4, 2005, 15:54:13  
Job time : 268.571 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 1004.86 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-4

Perfect score: 250

Sequence: 1 aggcctctatgatctacta.....agttcacatctgtggtcagt 250

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sta.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	250	100.0	250	6	AX343331 Sequence
2	250	100.0	675	9	M28223 Human JE ge
3	250	100.0	681	6	AX774779 Sequence
4	250	100.0	681	9	M30816 Human monoc
5	250	100.0	3221	9	AY357296 Homo sapi
6	250	100.0	3227	9	D26087 Human gene
7	250	100.0	11793	9	Y18933 Homo sapien
8	248.4	99.4	9174	9	AF519531 Homo sapi
9	248.4	99.4	147416	9	AC005549 Homo sapi
10	237.4	95.0	2776	9	M37719 Human monoc
11	162.2	64.9	5926	6	AX346555 Sequence
12	160.4	64.2	5926	6	AX346554 Sequence
13	51	20.4	3405	4	BOVMCP1X
14	39	15.6	279170	2	AC127645 Rattus no
15	37.6	15.0	145807	9	AC008692 Homo sapi
16	37.6	15.0	163788	2	AC018996 Homo sapi
17	37.6	15.0	174197	2	AC091837 Homo sapi
18	37.2	14.9	164138	2	AC109432 Rattus no
19	37.2	14.9	246991	2	AC094569 Rattus no

20	37.2	14.9	247021	2	AC094393	AC094393 Rattus no
21	37.2	14.9	334320	2	AC113256	AC113256 Rattus no
22	36.4	14.6	151899	2	AC111054	AC111054 Mus muscu
23	36	14.4	205533	10	AC118224	AC118224 Mus muscu
24	36	14.4	230183	2	AC113046	AC113046 Mus muscu
25	35.8	14.3	165	6	CQ467046	CQ467046 Sequence
26	35.4	14.2	235111	2	AC094776	AC094776 Rattus no
27	35.4	14.2	322783	2	AC133423	AC133423 Rattus no
28	35.2	14.1	201932	5	BX897729	BX897729 Zebrafish
29	35.2	14.1	205048	2	CR589944	CR589944 Danio rer
30	35.2	14.1	211382	9	AP001362	AP001362 Homo sapi
31	35.2	14.1	218490	10	AC113063	AC113063 Mus muscu
32	35.2	14.1	242467	2	BX927324	BX927324 Danio rer
33	35.2	14.1	245083	2	AC096203	AC096203 Rattus no
34	35.2	14.1	245900	2	AC097727	AC097727 Rattus no
35	35	14.0	158702	9	AC104472	AC104472 Homo sapi
36	35	14.0	187983	2	AC026210	AC026210 Homo sapi
37	34.8	13.9	148692	2	AC150159	AC150159 Gallus ga
38	34.8	13.9	190305	2	AC150054	AC150054 Gallus ga
39	34.8	13.9	200197	2	AC150055	AC150055 Gallus ga
40	34.6	13.8	168186	9	AC117385	AC117385 Homo sapi
41	34.6	13.8	171081	2	AC023402	AC023402 Homo sapi
42	34.6	13.8	184355	9	CR381572	CR381572 Human DNA
43	34.6	13.8	190199	2	AC120897	AC120897 Rattus no
44	34.6	13.8	312640	2	AC132785	AC132785 Rattus no
45	34.4	13.8	644	11	BV051306	BV051306 S212P6022

## ALIGNMENTS

RESULT 1	AX343331	Sequence 4 from Patent EP1170372.	250 bp	DNA	linear	PAT 01-FEB-2002
LOCUS	AX343331					
DEFINITION	AX343331					
ACCESSION	AX343331.1	GI:18491681				
VERSION						
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1					
AUTHORS	Roesl, F., Soto, U., Coy, J., Finzer, P., Delius, H., Poustka, A., zur Hausen, H. and Patzelt, A.					
TITLE	Regulatory sequences of the human mcp-1 gene					
JOURNAL	Patent: EP 1170372-A 4 09-JAN-2002;					
(DE)	DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS					

FEATURES	Location/Qualifiers
source	1..250
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

ORIGIN	Query Match	100.0%	Score 250;	DB 6;	Length 250;
	Best Local Similarity	100.0%	Pred. No. 8.4e-70;		
	Matches 250;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGCTTCTATGATGCTACTATTCTGCATTTCGAATGAGCAAAATGGATTTAATGCAATGTCATGCA 60			
Db	1	AGGCTTCTATGATGCTACTATTCTGCATTTCGAATGAGCAAAATGGATTTAATGCAATGTCATGCA 60			
Qy	61	GGGAGCCGGCCAAAGCTTGAGAGCTCCTTCCTGGCTGGGAGGCCCTTGGAAATGTGGCCT 120			
Db	61	GGGAGCCGGCCAAAGCTTGAGAGCTCCTTCCTGGCTGGGAGGCCCTTGGAAATGTGGCCT 120			
Qy	121	GAAGTAAGCTGGGAGCGGAGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCTTCCCTTTT 180			
Db	121	GAAGTAAGCTGGGAGCGGAGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCTTCCCTTTT 180			
Qy	181	TCTGAGTTTTTGGCTTTCACAGAAAGCAGATCCTTTAAAAATAACCCCTCTTAGTTTCACATC 240			

Tue Aug 9 17:01:26 2005

```
Db 181 TCTGAGTTTCGCTTCACAGAGCAATCCCTTAAAAATAACCCCTCTTAGTTCACATC 240
    |||
Qy 241 TGTGTCAGT 250
    |||
Db 241 TGTGTCAGT 250

RESULT 2
HUMSECP1
LOCUS Human JE gene encoding a monocyte secretory protein, exon 1.
DEFINITION M28223 M26035
ACCESSION M28223.1 GI:338003
VERSION M28223.1 GI:338003
KEYWORDS secretory protein.
SEGMENT 1 of 3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 675)
AUTHORS Rollins B.J., Stier, P., Ernst, T. and Wong, G.G.
TITLE The human homolog of the JE gene encodes a monocyte secretory
protein
JOURNAL Mol. Cell. Biol. 9 (11), 4687-4695 (1989)
MEDLINE 90097880
PUBMED 2513477
COMMENT Original source text: Human lung fibroblast WI38 cell line, DNA,
clones hJE-34 and lambda-hJE-7.
[1] sites for [Unpublished (1989) Dana-Farber Cancer Inst.,
Boston, Ma., 02115].
Draft entry and computer readable sequence for [Unpublished (1989)
Dana-Farber Cancer Inst., Boston, Ma., 02115] kindly provided by
B.J.Rollins 17-JUL-1989.
FEATURES
    source
    1..675
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /cell_line="WI38"
    /cell_type="fibroblast"
    /tissue_type="lung"
    532..675
    /gene="JE"
    /number=1
    600..668
    /gene="JE"

exon
sig_peptide
    250
    280 TGTGTCAGT 289

ORIGIN
    Query Match 100.0%; Score 250; DB 9; Length 675;
    Best Local Similarity 100.0%; Pred. No. 9.3e-70; Indels 0; Gaps 0;
    Matches 250; Conservative 0; Mismatches 0;
Qy 1 AGGCTTCTATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATGTGCA 60
    |||
Db 40 AGGCTTCTATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATGTGCA 99
    |||
Qy 61 GGGAGCGGCGCAAAAGCTTGAGAGCTCTTCTGCTGGGAGGCCCTTGGAAATGGCCCT 120
    |||
Db 100 GGGAGCGGCGCAAAAGCTTGAGAGCTCTTCTGCTGGGAGGCCCTTGGAAATGGCCCT 159
    |||
Qy 121 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCGCTTCCCTTCCTTT 180
    |||
Db 160 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCGCTTCCCTTCCTTT 219
    |||
Qy 181 TCTGAGTTTTCGCTTCACAGAAAGCAGAAATCCCTTAAAAATAACCCCTCTTAGTTCACATC 240
    |||
Db 220 TCTGAGTTTTCGCTTCACAGAAAGCAGAAATCCCTTAAAAATAACCCCTCTTAGTTCACATC 279
    |||
Qy 241 TGTGTCAGT 250
    |||
Db 280 TGTGTCAGT 289
```

```
RESULT 3
AX774779
LOCUS Sequence 95 from Patent WO03038129.
DEFINITION AX774779
ACCESSION AX774779
VERSION AX774779.1 GI:32486295
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Raponi, M.
TITLE Methods for assessing and treating leukemia
JOURNAL Patent: WO 03038129-A 95 08-MAY-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
    source
    1..681
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
    Query Match 100.0%; Score 250; DB 6; Length 681;
    Best Local Similarity 100.0%; Pred. No. 9.3e-70; Indels 0; Gaps 0;
    Matches 250; Conservative 0; Mismatches 0;
Qy 1 AGGCTTCTATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATGTGCA 60
    |||
Db 40 AGGCTTCTATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATGTGCA 99
    |||
Qy 61 GGGAGCGGCGCAAAAGCTTGAGAGCTCTTCTGCTGGGAGGCCCTTGGAAATGGCCCT 120
    |||
Db 100 GGGAGCGGCGCAAAAGCTTGAGAGCTCTTCTGCTGGGAGGCCCTTGGAAATGGCCCT 159
    |||
Qy 121 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCGCTTCCCTTCCTTT 180
    |||
Db 160 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCGCTTCCCTTCCTTT 219
    |||
Qy 181 TCTGAGTTTTCGCTTCACAGAAAGCAGAAATCCCTTAAAAATAACCCCTCTTAGTTCACATC 240
    |||
Db 220 TCTGAGTTTTCGCTTCACAGAAAGCAGAAATCCCTTAAAAATAACCCCTCTTAGTTCACATC 279
    |||
Qy 241 TGTGTCAGT 250
    |||
Db 280 TGTGTCAGT 289

RESULT 4
HUMSPAL
LOCUS Human monocyte secretory protein (JE) gene, exon 1.
DEFINITION HUMSPAL
ACCESSION M30816
VERSION M30816.1 GI:188699
KEYWORDS cytokine; secretory protein.
SEGMENT 1 of 3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Rollins B.J., Stier, P., Ernst, T. and Wong, G.G.
TITLE The human homolog of the JE gene encodes a monocyte secretory
protein
JOURNAL Mol. Cell. Biol. 9 (11), 4687-4695 (1989)
MEDLINE 90097880
PUBMED 2513477
COMMENT Original source text: Human WI-38 embryo lung fibroblast DNA, clone
hJE-34, passage 14.
FEATURES
    source
    1..681
    /organism="Homo sapiens"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
prim_transcript 532..>681
/feature="MSP mRNA and introns"
exon
<600..675
/feature="monocyte secretory protein"
/number=1
intron
676..>681
/feature="MSP, intron A"

ORIGIN
Query Match 100.0%; Score 250; DB 9; Length 681;
Best Local Similarity 100.0%; Pred. No. 9.3e-70;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTTCTATGATGCTACTATTCTGCATTTGATGAGCAATGGATTAAATGATGTGCA 60
Db 40 AGCTTCTATGATGCTACTATTCTGCATTTGATGAGCAATGGATTAAATGATGTGCA 99

QY 61 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGGAGGCCCTTGAATGTGGCCT 120
Db 100 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGGAGGCCCTTGAATGTGGCCT 159

QY 121 GAAGGTAAGCTGGCAGGAGCCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTT 180
Db 160 GAAGGTAAGCTGGCAGGAGCCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTT 219

QY 181 TCTGCAGTTTTCCTTCCACAGAACGAGATCCTTAAATAACCCCTTTAGTTTCATC 240
Db 220 TCTGCAGTTTTCCTTCCACAGAACGAGATCCTTAAATAACCCCTTTAGTTTCATC 279

QY 241 TGTGTCAGT 250
Db 280 TGTGTCAGT 289

RESULT 5
AY357296 3221 bp DNA linear PRI 15-SEP-2003
LOCUS Homo sapiens MCP1 (MCP1) gene, promoter region and partial cds.
DEFINITION
ACCESSION AY357296.1 GI:34559719
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Nyquist,P.A. and Degraaba,T.J.
TITLE Two New Single Nucleotide Polymorphisms (SNPs) in the MCP-1
Promoter in Patients with Carotid Atherosclerosis: Transcriptional
Induction and New Protein Binding Sites
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3221)
AUTHORS Nyquist,P.A. and Degraaba,T.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2003) Medicine, Inova Fairfax, 3300 Gallows Road,
Falls Church, VA 22402-3300, USA
FEATURES
source
1..3221
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
225..>3221
/feature="MCP1"
/notes="HSMCR30; MCP-1"
225..2746
/feature="MCP1"
/feature="MCP1"
225
/feature="MCP1"
/feature="SNP"
/replace="c"
668
variation
```

Tue Aug 9 17:01:26 2005

AUTHORS Shyy, Y.J., Li, Y.S. and Kolattukudy, P.E.  
 TITLE Structure of human monocyte chemotactic protein gene and its regulation by TPA  
 JOURNAL Biochem. Biophys. Res. Commun. 169 (2), 346-351 (1990)  
 MEDLINE 90290466  
 PUBMED 2357211  
 REFERENCE 3 (bases 1 to 3227)  
 AUTHORS Ueda, A., Okuda, K., Ohno, S., Shirai, A., Igarashi, T., Matsunaga, K., Fukushima, J., Kawamori, S., Ishigatsubo, Y. and Okubo, T.  
 TITLE NF-kappa B and Sp1 regulate transcription of the human monocyte chemoattractant protein-1 gene  
 JOURNAL J. Immunol. 153 (5), 2052-2063 (1994)  
 MEDLINE 94327939  
 PUBMED 8051410  
 REFERENCE 4 (bases 1 to 3227)  
 AUTHORS Ueda, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-DEC-1993) Atsuhisa Ueda, Yokohama City University School of Medicine, First Department of Internal Medicine; 3-9 Fukuura, Kanazawa-ku, Yokohama 236, Japan (Tel:045-787-2630, Fax:045-786-3444)  
 COMMENT Submitted (06-Dec-1993) to DDBJ by: Atsuhisa Ueda  
 First Department of Internal Medicine  
 Yokohama City University School of Medicine  
 3-9 Fukuura, Kanazawa-ku  
 Yokohama 236  
 Japan  
 Phone: 045-787-2630  
 Fax: 045-786-3444.  
 Location/Qualifiers  
 1. .3227  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="placenta"  
 /clone\_lib="EMBL3 SP6/T7"  
 400. .407  
 /note="kappa B"  
 2948. .2953  
 2975. .2979  
 3012. .3147  
 /number=1  
 3072. .>3147  
 /codon\_start=1  
 /product="monocyte chemoattractant protein-1"  
 /protein\_id="BAA05080.1"  
 /db\_xref="GI:992947"  
 /translation="MKVSAALLCLLLIAATFIPQGLAQP"  
 3148. .>3227  
 /number=1  
 enhancer  
 GC signal  
 TATA\_signal  
 exon  
 CDS  
 intron  
 ORIGIN  
 Query Match 100.0%; Score 250; DB 9; Length 3227;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-69;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGCTTCTATGATGCTACTATTCGATTTGCAATTCGAATGAGCAATGATTAATGCAATTCGA 60  
 DB 2512 AGGCTTCTATGATGCTACTATTCGATTTGCAATTCGAATGAGCAATGATTAATGCAATTCGA 2571  
 QY 61 GGGAGCGCGCCAAAGCTTGAGAGCTTCCTTCCTGGCTGGGAGGCCCTTTGGAATGTGGCCT 120  
 DB 2572 GGGAGCGCGCCAAAGCTTGAGAGCTTCCTTCCTGGCTGGGAGGCCCTTTGGAATGTGGCCT 2631  
 QY 121 GAAGGTAAAGCTGGCAGCGAGCGCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCCTTT 180  
 DB 2632 GAAGGTAAAGCTGGCAGCGAGCGCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCCTTT 2691  
 QY 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTTAAAAATAACCCCTCTTAGTTTCACATC 240  
 DB 2692 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTTAAAAATAACCCCTCTTAGTTTCACATC 2751

QY 241 TGTGTCAGT 250  
 DB 2752 TGTGTCAGT 2761  
 RESULT 7  
 HSY18933  
 LOCUS Homo sapiens MCP-1 gene and enhancer region. 11793 bp DNA linear PRI 18-OCT-2000  
 DEFINITION Y18933  
 ACCESSION Y18933  
 VERSION Y18933.1 GI:10933860  
 KEYWORDS MCP-1 gene; monocyte chemoattractant protein-1.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Finzer, P., Soto, J., Delius, H., Patzelt, A., Coy, J.F., Poustka, A., zur Hausen, H. and Rosl, F.  
 TITLE Differential transcriptional regulation of the monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and non-tumorigenic HPV 18 positive cells: the role of the chromatin structure and AP-1 composition  
 JOURNAL Oncogene 19 (29), 3235-3244 (2000)  
 MEDLINE 20374005  
 PUBMED 10918580  
 REFERENCE 2 (bases 1 to 11793)  
 AUTHORS Roesl, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-APR-1999) F. Roesl, Applied Tumor Virology, DKFZ (German Cancer Research Institute), Im Neuenheimer Feld 506, 69120 Heidelberg, FRG  
 FEATURES  
 Location/Qualifiers  
 1. .11793  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 4906. .9073  
 /gene="MCP-1"  
 4906. .4915  
 /gene="MCP-1"  
 4935. .4944  
 /gene="MCP-1"  
 7513. .7517  
 /gene="MCP-1"  
 join(<7550. .7685,8468. .8585,8968. .>9073)  
 /gene="MCP-1"  
 7550. .7685  
 /gene="MCP-1"  
 /number=1  
 join(7610. .7685,8468. .8585,8968. .9073)  
 /gene="MCP-1"  
 /codon\_start=1  
 /product="monocyte chemoattractant protein-1"  
 /protein\_id="CAC14049.1"  
 /db\_xref="GI:10933861"  
 /db\_xref="GOA:P13500"  
 /db\_xref="UniProt/Swiss-Prot:P13500"  
 /translation="MKVSAALLCLLLIAATFIPQGLAQPDAINAPVTCYAFNFKISVORLASVRRITSSKCPKEAIVFKTIVAKEICADPKQKWQDSMDHLDKQTQPKT"  
 7686. .8467  
 /gene="MCP-1"  
 /number=1  
 8468. .8585  
 /gene="MCP-1"  
 /number=2  
 8586. .8967  
 /gene="MCP-1"  
 /number=2  
 8968. .>9073  
 /gene="MCP-1"  
 /number=3  
 intron  
 exon  
 intron  
 exon

## ORIGIN

Query Match 100.0%; Score 250; DB 9; Length 11793;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-69;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTAATGCATTGTCA 60  
 |||||  
 Db 7050 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTAATGCATTGTCA 7109  
 |||||

QY 61 GGGAGCGGCCCAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGCGCCT 120  
 |||||  
 Db 7110 GGGAGCGGCCCAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGCGCCT 7169  
 |||||

QY 121 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 180  
 |||||  
 Db 7170 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 7229  
 |||||

QY 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAATAACCTCTTAGTTCACATC 240  
 |||||  
 Db 7230 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAATAACCTCTTAGTTCACATC 7289  
 |||||

QY 241 TGTGTCAGT 250  
 |||||  
 Db 7290 TGTGTCAGT 7299  
 |||||

## RESULT 8

AF519531  
 LOCUS 9174 bp DNA linear PRI 09-JUL-2002  
 DEFINITION Homo sapiens small inducible cytokine A2 (monocyte chemotactic protein 1) (SCYA2) gene, complete cds.

ACCESSION AF519531  
 VERSION AF519531.1 GI:21435976  
 KEYWORDS

SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Rieder,M.J., Armel,T.Z., Carrington,D.P., Ozuna,M., Kuldane,K.S.A.,  
 Rajkumar,N., Toth,E.J., Yi,Q. and Nickerson,D.A.  
 TITLE Submitted (06-JUN-2002) Genome Sciences, University of Washington,  
 1705 NE Pacific, Seattle, WA 98195, USA  
 JOURNAL Direct Submission  
 COMMENT To cite this work please use: SeattleSNPs. NHLBI HL66682 Program  
 for Genomic Applications, UW-FHCRC, Seattle, WA (URL:  
 http://pga.gs.washington.edu).

## FEATURES

source  
 1..9174  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

variation  
 107  
 /frequency="0.17"  
 /replace="g"

variation  
 291  
 /frequency="0.03"  
 /replace="g"

variation  
 305  
 /frequency="0.15"  
 /replace="t"

variation  
 334  
 /frequency="0.01"  
 /replace="g"

variation  
 604  
 /frequency="0.16"  
 /replace="t"

variation  
 609  
 /frequency="0.02"  
 /replace="a"

repeat\_region  
 614..720  
 /rpt\_family="MIR"

variation  
 678  
 /rpt\_type="dispersed"  
 /frequency="0.14"  
 /replace="g"

variation  
 1051  
 /frequency="0.01"  
 /replace="a"

variation  
 1054  
 /frequency="0.01"  
 /replace="c"

variation  
 1072  
 /frequency="0.03"  
 /replace="t"

variation  
 1280  
 /frequency="0.23"  
 /replace="a"

variation  
 1537  
 /frequency="0.15"  
 /replace="g"

variation  
 1548  
 /frequency="0.08"  
 /replace="a"

variation  
 1563  
 /frequency="0.04"  
 /replace="g"

variation  
 1707  
 /frequency="0.02"  
 /replace="c"

repeat\_region  
 1718..1830  
 /rpt\_family="MIR"

variation  
 1729  
 /rpt\_type="dispersed"  
 /frequency="0.01"  
 /replace="c"

variation  
 1980  
 /frequency="0.13"  
 /replace="t"

variation  
 1984  
 /frequency="0.02"  
 /replace="g"

variation  
 2307  
 /frequency="0.04"  
 /replace="g"

variation  
 2391  
 /frequency="0.01"  
 /replace="a"

variation  
 2569  
 /frequency="0.08"  
 /replace="t"

variation  
 3001  
 /frequency="0.03"  
 /replace="a"

variation  
 3190  
 /frequency="0.14"  
 /replace="c"

repeat\_region  
 3228..3445  
 /rpt\_family="MIR"

variation  
 3719  
 /rpt\_type="dispersed"  
 /frequency="0.01"  
 /replace="a"

variation  
 3756  
 /frequency="0.25"  
 /replace="c"

gene  
 4065..5581  
 /gene="SCYA2"

mRNA  
 join(4065..4193,4990..5107,5490..5581)  
 /product="small inducible cytokine A2 (monocyte chemotactic protein 1)"  
 /gene="SCYA2"

CDS  
 join(4118..4193,4990..5107,5490..5595)  
 /gene="SCYA2"  
 /codon\_start=1

```

/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
protein id="AAM54046.1"
/db_xref="GI:21435977"
/translation="MKVSAALLCLLLIAATFIPOGLAOPDAINAPVTCCYNFTNRKIS
VORLASVRRITSSKCPKEAVIFKTVIAKEICADPKQKWQDSMDHLDKQTQPKT"
4296..4409
/repeat_region
  /rpt_family="MIR"
  /rpt_type="dispersed"
  /rpt_type="dispersed"
  /gene="SCYA2"
  /note="Region not scanned for variation"
  4748
  /gene="SCYA2"
  /frequency="0.26"
  /replace="n"
  4777..4814
  /repeat_region
    /rpt_family="MIR"
    /rpt_type="dispersed"
    4881
    /gene="SCYA2"
    /frequency="0.14"
    /replace="g"
    5018
    /gene="SCYA2"
    /frequency="0.47"
    /replace="c"
    5132
    /gene="SCYA2"
    /frequency="0.01"
    /replace="t"
    5660
    /gene="SCYA2"
    /frequency="0.18"
    /replace="t"
    6184
    /frequency="0.01"
    /replace="c"
    6518
    /frequency="0.01"
    /replace="t"
    6531
    /frequency="0.05"
    /replace="t"
    6837..6916
    /rpt_family="L2"
    /rpt_type="dispersed"
    7436
    /frequency="0.07"
    /replace="t"
    7605
    /frequency="0.03"
    /replace="a"
    7614
    /frequency="0.01"
    /replace="a"
    7659
    /frequency="0.03"
    /replace="a"
    7710
    /frequency="0.05"
    /replace="a"
    7843
    /frequency="0.31"
    /replace="c"
    8072
    /frequency="0.01"
    /replace="t"
    8118..8271
    /rpt_family="MaLR"
    /rpt_type="dispersed"
    8272..8734
    /rpt_family="MaLR"
  
```

```

/variation
  /rpt_type="dispersed"
  8597
  /frequency="0.07"
  /replace="c"
  8735..9033
  /repeat_region
    /rpt_family="Alu"
    /rpt_type="dispersed"
    8775
    /frequency="0.07"
    /replace="a"
  
```

Query Match 99.4%; Score 248.4; DB 9; Length 9174;  
 Best Local Similarity 99.6%; Pred. No. 3.9e-69;  
 Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGCAATGTCA 60
  |||||
Db 3558 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGCAATGTCA 3617
  |||||
QY 61 GGGAGCGCGCCAAAGCTTGAGAGCTCTTCCCTGGCTGGGAGGCCCTTGGAAATGTGGCCT 120
  |||||
Db 3618 GGGAGCGCGCCAAAGCTTGAGAGCTCTTCCCTGGCTGGGAGGCCCTTGGAAATGTGGCCT 3677
  |||||
QY 121 GAAGGTAAGCTGGCAGCGACCTGACATGCTTTTCATCTAGTTTCTCGCTTCTCTCTCTTT 180
  |||||
Db 3678 GAAGGTAAGCTGGCAGCGACCTGACATGCTTTTCATCTAGTTTCTCGCTTCTCTCTCTTT 3737
  |||||
QY 181 TCTGCAGTTTTCGCTTACAGAAAGCAGATCCTTAAATAATACCCCTCTTCTCTCTCTCT 240
  |||||
Db 3738 TCTGCAGTTTTCGCTTACAGAAAGCAGATCCTTAAATAATACCCCTCTTCTCTCTCTCT 3797
  |||||
QY 241 TGTGTCAGT 250
  |||||
Db 3798 TGTGTCAGT 3807
  |||||

```

RESULT 9  
 AC005549/c 147416 bp DNA linear PRI 22-SEP-1998  
 LOCUS Homo sapiens chromosome 17, clone hRPK.215\_E13, complete sequence.  
 DEFINITION AC005549  
 ACCESSION AC005549.1 GI:3598724  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 1 (bases 1 to 147416)  
 2 (bases 1 to 147416)  
 Unpublished  
 1 (bases 1 to 147416)  
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
 Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J.,  
 Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M.,  
 Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E.,  
 Devon,K., Dewar,K., Donelan,L., Ferreira,P., Fitzhugh,W.,  
 Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G.,  
 Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C.,  
 Jacotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J.,  
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
 Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,  
 Nafz,R., Naylor,J., Niloff,N., O'Connor,T., O'Donnell,P.,  
 Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,  
 Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C.,  
 Subramanian,A., Tesfaye,S., Tichovolsky,N., Torrella-Miller,I.,  
 Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D.,  
 Ye,W.J., Zhao,J. and Zody,M.  
 Direct Submission  
 Submitted (27-AUG-1998) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 147416)  
 Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Corliss, D., Depayre, E., Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraghty, K., Grant, G., Hagos, B., Headford, A., Herena, L., Horton, L., Howland, J.C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczeky, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W.J., Zhao, J. and Zody, M.

# Direct Submission

Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 147416)

## REFERENCE AUTHORS

Birren, B., Pasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Corliss, D., Depayre, E., Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraghty, K., Grant, G., Hagos, B., Headford, A., Herena, L., Horton, L., Howland, J.C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczeky, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W.J., Zhao, J. and Zody, M.

# Direct Submission

Submitted (22-SEP-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 13, 1998 this sequence version replaced gi:3581743.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only the first 147416 bases of this clone are being submitted.

Bases 145417-155040 overlap accession number AC004147 (WICGR Project L228). The first 2Kb of the overlapping region are submitted to confirm overlap.

## FEATURES

### Source

1. 147416  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RRPK.215 E 13"  
/clone\_lib="RPCI-11 human BAC library"  
2. .65  
/rpt\_family="BUR1"  
490. .695  
/rpt\_family="LIME2"  
1291. .1390  
/rpt\_family="MIR"  
2980. .3131  
/rpt\_family="MER5A"  
complement(3147. .3241)  
/rpt\_family="MIR"  
complement(4228. .4469)  
/rpt\_family="MLT1"  
complement(5430. .5549)  
/rpt\_family="FLAM\_C"  
complement(5941. .6051)  
/rpt\_family="MER81"  
6131. .6250

repeat\_region  
6406. .6455  
/rpt\_family="AT\_rich"  
6574. .6600  
/rpt\_family="MLT1D"  
6601. .6880  
/rpt\_family="AluJb"  
6881. .6935  
/rpt\_family="MLT1D"  
7119. .7360  
/rpt\_family="MLT1D"  
7337. .7636  
/rpt\_family="MLT1D"  
7689. .7888  
/rpt\_family="LIME3"  
8023. .8048  
/rpt\_family="AT\_rich"  
8871. .8986  
/rpt\_family="MIR"  
10109. .10258  
/rpt\_family="MER5A"  
complement(11031. .11072)  
/rpt\_family="MIR"  
11567. .11836  
/rpt\_family="L2"  
12501. .12981  
/rpt\_family="MER9"  
complement(12983. .13257)  
/rpt\_family="AluJb"  
15248. .15364  
/rpt\_family="MIR"  
complement(16483. .16653)  
/rpt\_family="MIR"  
16830. .16949  
/rpt\_family="MER5A"  
17203. .17380  
/rpt\_family="LTR16A"  
17393. .18306  
/rpt\_family="HERVL"  
18315. .24307  
/rpt\_family="LIP2"  
24309. .24335  
/rpt\_family="AT\_rich"  
24513. .25245  
/rpt\_family="HERVL"  
25613. .25758  
/rpt\_family="HERVL"  
26135. .26228  
/rpt\_family="LTR16A"  
complement(27814. .27950)  
/rpt\_family="MER94"  
28930. .28968  
/rpt\_family="AT\_rich"  
29960. .30341  
/rpt\_family="MLT1"  
32364. .32581  
/rpt\_family="MIR"  
32601. .32902  
/rpt\_family="AluSx"  
complement(33593. .39359)  
/rpt\_family="LIP2"  
39362. .39414  
/rpt\_family="A-rich"  
complement(39920. .40047)  
/rpt\_family="MIR"  
40112. .40151  
/rpt\_family="AT\_rich"  
complement(40152. .40442)  
/rpt\_family="AluJo"  
complement(40810. .40898)  
/rpt\_family="MLT1D"  
41293. .41469  
/rpt\_family="LIP2"

Tue Aug 9 17:01:26 2005

```
repeat_region complement(41846..41956)
repeat_region /rpt_family="purine-rich"
repeat_region complement(42486..42569)
repeat_region /rpt_family="MLT11"
repeat_region 42975..43098
repeat_region /rpt_family="MERSA"
repeat_region 43554..43983
repeat_region /rpt_family="LTR24"
repeat_region 44781..44811
repeat_region /rpt_family="AT-rich"
repeat_region 46184..46296
repeat_region /rpt_family="MIR"
repeat_region complement(47075..47140)
repeat_region /rpt_family="MLT1G"
repeat_region complement(47266..47307)
repeat_region /rpt_family="(CA)n"
repeat_region complement(47363..47442)
repeat_region /rpt_family="MLT1G"
repeat_region complement(48165..48209)

Query Match 99.4%; Score 248.4; DB 9; Length 147416;
Best Local Similarity 99.6%; Pred. No. 5.1e-69;
Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGCTACTACTCTGCAATTTGAATGAGCAAAATGGATTTAATGCAATGTCA 60
Db 61977 AGGCTTCTATGCTACTACTCTGCAATTTGAATGAGCAAAATGGATTTAATGCAATGTCA 61918

QY 61 GGGAGCGGGCCAAAGCTTGAAGCTCTTCTGGCTGGGAGGCCCTTGGAAATGGGCTT 120
Db 61917 GGGAGCGGGCCAAAGCTTGAAGCTCTTCTGGCTGGGAGGCCCTTGGAAATGGGCTT 61858

QY 121 GAAGTTAAGCTGGAGGAGCGCTGACATCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 180
Db 61857 GAAGTTAAGCTGGAGGAGCGCTGACATCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 61798

QY 181 TCTGAGTTTCGGCTTTCACAGAAAGCAGAAATCCTTAAATAAACCCTCTTAGTTACATC 240
Db 61797 TCTGAGTTTCGGCTTTCACAGAAAGCAGAAATCCTTAAATAAACCCTCTTAGTTACATC 61738

QY 241 TGTGTCAGT 250
Db 61737 TGTGTCAGT 61728

RESULT 10
HUMMCHEMP 2776 bp DNA linear PRI 13-MAY-1994
LOCUS Human monocyte chemotactic protein gene, complete cds.
DEFINITION M37719
ACCESSION M37719.1 GI:187447
VERSION monocyte chemotactic protein.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2776)
AUTHORS Shyu,Y.J., Li,Y.S. and Kolattukudy,P.E.
TITLE Structure of human monocyte chemotactic protein gene and its regulation by TPA
JOURNAL Biochem. Biophys. Res. Commun. 169 (2), 346-351 (1990)
MEDLINE 90290466
PUBMED 2357211
COMMENT source text: Human DNA.
FEATURES Location/Qualifiers
source 1..2776
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
598..2080
/gene="SCYA2"
join(598..673,1472..1589,1975..2080)
/gene="SCYA2"
```

```
/note="monocyte chemotactic protein"
/codon_start=1
/protein_id="AAA18102.1"
/db_xref="GI:487124"
/translation="MKVSAALLCLLLIAATFIPOGLAQPDAINAPVTCVNFTRKIS
VORLASVRRITSSKCPKEAVIFKTIKAECADPKQWVQDSMDHLDKQTQTPKT"
<598..673
/gene="SCYA2"
/note="monocyte chemotactic protein"
/number=1
598..673
/gene="SCYA2"
/note="monocyte chemotactic protein"
/number=1
674..1471
/gene="SCYA2"
/note="monocyte chemotactic protein intron A"
1472..1589
/gene="SCYA2"
/note="monocyte chemotactic protein"
/number=2
1590..1974
/gene="SCYA2"
/note="monocyte chemotactic protein intron B"
1975..2080
/gene="SCYA2"
/note="monocyte chemotactic protein"
/number=3
1975..2080
/gene="SCYA2"
/note="monocyte chemotactic protein"
/number=3

ORIGIN
Query Match 95.0%; Score 237.4; DB 9; Length 2776;
Best Local Similarity 99.2%; Pred. No. 1.3e-65;
Matches 249; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AGGCTTCTATGCTACTACTCTGCAATTTGAATGAGCAAAATGGATTTAATGCAATGTCA 60
Db 35 AGGCTTCTATGCTACTACTCTGCAATTTGAATGAGCAAAATGGATTTAATGCAATGTCA 94

QY 61 GGGAGCGGGCCAAAGCTTGAAGCTCTTCTGGCTGGGAGGCCCTTGGAAATGGGCTT 120
Db 95 GGGAGCGGGCCAAAGCTTGAAGCTCTTCTGGCTGGGAGGCCCTTGGAAATGGGCTT 154

QY 121 GAAGTTAAGCTGGAGGAGCGCTGACATCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 180
Db 155 GAAGTTAAGCTGGAGGAGCGCTGACATCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 214

QY 181 T-CTGAGTTTCGGCTTTCACAGAAAGCAGAAATCCTTAAATAAACCCTCTTAGTTACAT 239
Db 215 TCTGCAAGTTTCGGCTTTCAGAGAAAGCAGAAATCCTTAAATAAACCCTCTTAGTTACAT 274

QY 240 CTGTCGTGTCAGT 250
Db 275 CTGTCGTGTCAGT 285

RESULT 11
AX346555/c
LOCUS AX346555
DEFINITION Sequence 1626 from Patent WO0200928.
ACCESSION AX346555
VERSION AX346555.1 GI:18494441
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1626 03-JAN-2002;
```

Epigenomics AG (DE)  
Location/Qualifiers  
1. .5926  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 64.9%; Score 162.2; DB 6; Length 5926;  
Best Local Similarity 78.5%; Pred. No. 3.6e-41;  
Matches 194; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 4 CTTCTATGATGCTACTTCTGCAATTTGAATGAGCAAAATGGATTAAATGCAATTTGTCAGGG 63  
|||||  
Db 1430 CTTCTATGATGCTACTTCTGCAATTTGAATGAGCAAAATGGATTAAATGCAATTTGTCAGGG 63  
|||||

QY 64 AGCGGCGCAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCGCCCTTGGAAATGTGGCCTGAA 123  
|||||  
Db 1370 AACCGACCAAACTTAAAACTCTTCCTAACTTAAAAAACCCCTTAAATAATACCTAAA 1311  
|||||

QY 124 GGTAACTGGCAGCGAGCTGACATGCTTTTCATCTAGTTTTCCTCGCTTCCTTCCTTCCT 183  
|||||  
Db 1310 AATAACTAACACGAACTTAACTATTTTCATCTAGTTTTCCTCGCTTCCTTCCTTCCT 1251  
|||||

QY 184 GCAGTTTTCGCTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTCACATCTCT 243  
|||||  
Db 1250 ACAATTTTCGCTTCAAAAACCAAAATCCTTAAATAATACCTCTTAGTTCACATCTAT 1191  
|||||

QY 244 GGTCACT 250  
|||||  
Db 1190 AATCAAT 1184  
|||||

RESULT 12  
AX346554  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

AX346554 5926 bp DNA linear PAT 01-FEB-2002  
Sequence from Patent WO0200928.  
AX346554  
AX346554.1 GI:18494440  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
1  
Olek A., Piepenbrock C. and Berlin K.  
Diagnosis of diseases associated with the immune system  
Patent: WO 0200928-A 1625 03-JAN-2002;  
Epigenomics AG (DE)  
Location/Qualifiers  
1. .5926  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 64.2%; Score 160.4; DB 6; Length 5926;  
Best Local Similarity 77.5%; Pred. No. 1.4e-40;  
Matches 194; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGATGCTACTTCTGCAATTTGAATGAGCAAAATGGATTAAATGCAATTTGTC 60  
|||||  
Db 4494 AGGTTTATGATGTTATTTATTTGTTATTTGAATGAGTAAATGGATTAAATGTTATTTGTA 4553  
|||||

QY 61 GGGAGCGGCGCAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCGCCCTTGGAAATGTGGCT 120  
|||||  
Db 4554 GGGAGTCGGTTAAAGTTTCAGAGTTTTCCTGGTGGGAGGTTTTCCTGGGAGGTTTTCCTGG 4613  
|||||

QY 121 GAAGGTAAGCTGGCAGCGAGCTGACATGCTTTTCATCTAGTTTTCCTCGCTTCCTTCCTTT 180  
|||||  
Db 4614 GAAGGTAAGCTGGCAGCGAGTTGATGATGTTTATTTATTTAGTTTTCCTCGTTTTCCTTTT 4673  
|||||

QY 181 TCTGCAGTTTTCGCTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTCACATC 240  
|||||  
Db 4674 TTTGTAGTTTTCGTTTAGAGAAAGTAGAATTTTAAATAATATTTTTCAGTTTATATT 4733  
|||||

QY 241 TGTGTCAGT 250  
|||||  
Db 4734 TGTGTTAGT 4743  
|||||

RESULT 13  
BOVMCP1X  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
FEATURES  
source

BOVMCP1X 3405 bp DNA linear MAM 17-JAN-1995  
Bovine monocyte chemoattractant protein-1 (MCP-1) gene exons 1-3,  
complete cds.  
L32659  
L32659.1 GI:624393  
monocyte chemoattractant protein-1.  
Bos taurus (cow)  
Bos taurus  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
1 (bases 1 to 3405)  
Wempe, F., Kuhlmann, J.K. and Scheit, K.H.  
Characterization of the bovine monocyte chemoattractant protein-1  
Gene  
Biochem. Biophys. Res. Commun. 202 (3), 1272-1279 (1994)  
94338337  
8060303  
Original  
source text: Bos taurus male DNA.  
Location/Qualifiers  
1. .3405  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9913"  
/sex="male"  
/cell\_type="monocyte"  
1310..1315  
/gene="MCP-1"  
/note="putative"  
1342..1478  
/gene="MCP-1"  
/number=1  
join(1403..1478,2227..2344,2695..2800)  
/gene="MCP-1"  
join(1403..1478,2227..2344,2695..2800)  
/gene="MCP-1"  
/codon\_start=1  
/product="monocyte chemoattractant protein-1"  
/protein\_id="AAA60956.1"  
/db\_xref="GI:624394"  
/translation="MKVSAALLCLLLTVAAFTSLVLAQPDAINSQVACCYTNSKKIS  
MQLMNYRRVTSSKCPKEAVIFKTLGKELCADPKQKRWQDSINLYLNKQTPKP"

TATA\_signal  
exon  
gene  
CDS  
intron  
exon  
intron  
exon  
polyA\_signal  
ORIGIN

source text: Bos taurus male DNA.  
Location/Qualifiers  
1. .3405  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9913"  
/sex="male"  
/cell\_type="monocyte"  
1310..1315  
/gene="MCP-1"  
/note="putative"  
1342..1478  
/gene="MCP-1"  
/number=1  
join(1403..1478,2227..2344,2695..2800)  
/gene="MCP-1"  
join(1403..1478,2227..2344,2695..2800)  
/gene="MCP-1"  
/codon\_start=1  
/product="monocyte chemoattractant protein-1"  
/protein\_id="AAA60956.1"  
/db\_xref="GI:624394"  
/translation="MKVSAALLCLLLTVAAFTSLVLAQPDAINSQVACCYTNSKKIS  
MQLMNYRRVTSSKCPKEAVIFKTLGKELCADPKQKRWQDSINLYLNKQTPKP"

Query Match 20.4%; Score 51; DB 4; Length 3405;  
Best Local Similarity 72.1%; Pred. No. 4.4e-05;  
Matches 80; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY	1	AGGCTTCTATGATGCTACTATCTTCGATTTGAATGACCAATGGAATTAATGATTCATGTC	60
Db	684	AAGCTTCACTTTGTACCACTATCTACATTTTAAAGAGCAAAAGGATTTAATACGCTGTG	743
QY	61	GGGA-GCCGGCAAGCTTGAGAGCTCCTTCCTCGCTGGGAGGCCCTTCG	110
Db	744	GGGGTGCTGCCAGACTTCAGGACTCCTCTTGCTTGAGGTCCTCTGGG	794
RESULT 14			
AC127645			
LOCUS			
DEFINITION		Rattus norvegicus clone CH230-230H10, *** SEQUENCING IN PROGRESS	
ACCESSION		AC127645.3	
VERSION		GI:23195204	
KEYWORDS		HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.	
SOURCE		Rattus norvegicus (Norway rat)	
ORGANISM		Rattus norvegicus	
REFERENCE			
AUTHORS		1 (bases 1 to 279170) Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blunk, P., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Ceasar, H., Center, A., Cardenas, J., Carter, K., Cavazos, I., Chen, R., Chen, Y., Chen, Z., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Coyle, M., Cree, A., D'Souza, L., Cleveland, C., Cockrell, R., Cox, C., Coyler, L., De Anda, C., Dederich, D., Davila, M., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Delgado, O., Denson, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Draper, H., Dugan-Rocha, S., Dunn, A., Evans, C.A., Falle, T., Fan, G., Egan, A., Escotto, M., Eugene, C., Forbes, L., Foster, M., Foster, P., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Garza, M., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunnatne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, I., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, D., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarpunsaogon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Pua, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Sma, J., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Sma, J., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingle, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, V., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederthausen, A., Weis, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.	
TITLE			

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 279170)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 279170)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (08-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Sep 19, 2002 this sequence version replaced gi:21953804. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
	----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GRW1 Center clone name: CH230-230H10 ----- Summary Statistics Assembly program: Phrap; version 0.990329 Consensus quality: 146772 bases at least Q40 Consensus quality: 148340 bases at least Q30 Consensus quality: 149405 bases at least Q20 Estimated insert size: 164568; sum-of-contigs estimation Quality coverage: 3x in Q20 bases; sum-of-contigs estimation ----- * NOTE: Estimated insert size may differ from sequence length * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. * 1 241705: contig of 241705 bp in length * 241706 241805: gap of unknown length * 241806 279170: contig of 37365 bp in length.
FEATURES	Location/Qualifiers
source	1..279170 /organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:10116" /clone="CH230-230H10" 236659..238546 /notes="wgs_contig" 240512..241705 /note="wgs_contig"
misc_feature	
misc_feature	
ORIGIN	
Query Match	15.6%; Score 39; DB 2; Length 279170;
Best Local Similarity	58.0%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches	69; Conservative 0; Mismatches 50;
QY	120 TGAAGTAAAGTGGCAGCGAGCTGCATGCTTTCTATCTAGTTCTCGTTCCTTCCTT 179 

Db 48507 TGAATGGATCCAAATGAGCGCTGACAGACTTCCTTAGTCTCTCTCTATTTTGTG 48566

Qy 180 TTCTCGAGTTTTCGCTTTCAGAAAGCAGAACTCTTAAAAATAACCTCTTAGTTTCA 238

Db 48567 TTGTAGTTTTTAGATTTGAACAATCTGGATCAGAAATTTTGTCTGTGGTTTAAACA 48625

RESULT 15  
AC008692  
LOCUS  
AC008692 145807 bp DNA linear PRI 18-DEC-2001  
DEFINITION Homo sapiens chromosome 5 clone CTB-65N22, complete sequence.

Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 145807)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS

TITLE	Direct Submission	Unpublished
JOURNAL		

REFERENCE 2 (bases 1 to 145807)  
AUTHORS DOE Joint Genome Institute.

**JOURNAL**  
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek CA 94598, USA

REFERENCE  
3 (bases 1 to 145807)  
DOE Joint Genome Institute and Stanford Human Genome Center,  
AUTHORS

**JOURNAL TITLE**  
Direct Submission  
Submitted (18-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT  
On Dec 18, 2001 this sequence version replaced gi:9256030.  
Drift Sequence Produced by DOG Joint Genome Institute  
Drive, Walnut Creek, CA 94598, USA

Plant Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

[www-shgc.stanford.edu](http://www-shgc.stanford.edu)

Quality: Phrap Quality >=40 100% of Sequence;

[illegible]

```
source
1. .145807
```

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"
```

```
/mol_type="genomic DNA"  
/db_xref="taxon:9606"
```

```
/w_xref="taxon:9606"  
/chromosome="5"
```

## ORIGIN

Query Match 15.0%; Score 37.6; DB 9; Length 145807;

Best Local Similarity 53.4%; Pred. No. 1.4;

Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 88 TTCTTGGCTGGAGGCCCTTTGGAAATGTGGCCTGAAGTAAGCTGGCAGCGAGCCTGACA 147

— — — — —

D**b** 102628 TGCATGAGAGTGAGAGCTGATGGGGCGGGCCAGATGAAATGCAGAAAGCACTTTTAA 102687

[illegible]

QY  
148 TGCCTTCATC TAGTTCCTCGCTTCCTCCCTTCCTGAGTTTTCGCTTCACAGAAAGCA 207

db 102688 TACTTTTTTTGGATGTGTGTTGCTTATACCTATTTGTCTGTTGATGTTCTGTGAAAAATA 102747

Search completed: August 4, 2005, 11:16:50  
Job time : 1009.86 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 165.688 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-4

Perfect score: 250

Sequence: 1 aggcctctatgatgctacta.....agttcacatctgtggtcagt 250

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	250	100.0	250	6	ADH13941 Human mon
2	250	100.0	681	10	ADP84876 Farnesyl
3	250	100.0	3221	12	ADN12120 MCP1 gene
4	250	100.0	3221	12	ADO03802 Human Ccl
5	250	100.0	11793	12	ADO03803 Human Ccl
6	237.4	95.0	2775	6	ABK47979 Human sma
7	162.2	64.9	5926	6	ABL33653 Human imm
8	160.4	64.2	5926	6	ABL33652 Human imm
9	91.4	36.6	800	8	ACF64398 Human MCP
10	36.4	14.6	80275	12	ADQ97310 Mouse can
11	36.4	14.6	109559	13	ABD33157 Murine ca
12	35.8	14.3	165	6	ABL86846 Human ova
13	34	13.6	1500	5	AAS88382 DNA encod
14	33.8	13.5	31051	4	AAK73223 Human imm
15	33.6	13.4	673	6	ABK35434 Human cDN
16	33.2	13.3	227968	6	ABK83497 Human cDN
17	33.2	13.3	227968	12	ADQ18538 Human sof
18	33	13.2	141586	11	ABD20695 Human pul
19	32.8	13.1	23406	11	ACN44442 Human gen
20	32.8	13.1	260160	12	ADQ20017 Human sof

C 21	32.6	13.0	2092	6	ABK81833	Abk81833 DNA repre
C 22	32.4	13.0	2943	4	ABN86540	Abn86540 N. tabacu
C 23	32	12.8	469	13	ACN47169	Acn47169 Cotton pr
C 24	31.8	12.7	463	4	AAK64539	Aak64539 Human imm
C 25	31.8	12.7	495	13	ADO52904	Ado52904 Novel can
C 26	31.6	12.6	1118	5	ABV24663	Abv24663 Human pro
C 27	31.6	12.6	2000	8	ADA72185	Ada72185 Rice gene
C 28	31.6	12.6	3417	8	ADA48709	Ada48709 Rat metab
C 29	31.6	12.6	5879	12	ADJ75271	Adj75271 Marker ge
C 30	31.6	12.6	5879	12	ADJ75181	Adj75181 Marker ge
C 31	31.6	12.6	5879	13	ADR24834	Adr24834 Breast ca
C 32	31.6	12.6	5920	12	ADJ75182	Adj75182 Marker ge
C 33	31.6	12.6	6921	8	ABX76394	Abx76394 Lung canc
C 34	31.6	12.6	6921	8	ABX76140	Abx76140 Lung canc
C 35	31.6	12.6	6921	11	ADN39009	Adn39009 Cancer/an
C 36	31.6	12.6	7043	8	ABX76393	Abx76393 Lung canc
C 37	31.6	12.6	7043	8	ABX76141	Abx76141 Lung canc
C 38	31.6	12.6	7043	11	ADN39007	Adn39007 Cancer/an
C 39	31.2	12.5	662	4	AAK57161	Aak57161 Human imm
C 40	31.2	12.5	2118	4	AAK65201	Aak65201 Human imm
C 41	31.2	12.5	2118	4	AAK65200	Aak65200 Human imm
C 42	31.2	12.5	51615	11	ACN45162	Acn45162 Human gen
C 43	30.8	12.3	2285	4	ABL27938	Ab127938 Drosophil
C 44	30.8	12.3	4491	8	ACA30151	Ac30151 Prokaryot
C 45	30.6	12.2	1074	8	ACD05595	Acd05595 cDNA enco

ALIGNMENTS

RESULT 1

ADH13941  
ID ADH13941 standard; DNA; 250 BP.

AC ADH13941;

DT 11-MAR-2004 (first entry)

DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:4.  
KW ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
KW cervical carcinoma.

OS Homo sapiens.

XX

PN EP1170372-A1.

XX

PD 09-JAN-2002.

XX

PF 06-JUL-2000; 2000EP-00114560.

XX

PR 06-JUL-2000; 2000EP-00114560.

XX

PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX

PI Roel F, Soto U, Coy J, Finzer P, Delius H, Poustka A;

PI Zur Hausen H, Patzelt A;

XX

XX WPI; 2002-165895/22.

DR

PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,

PT useful in gene therapy, for treating atherosclerosis and cancer.

XX

PS Disclosure; SEQ ID NO 4; 30pp; English.

XX

CC The invention relates to a novel nucleic acid molecule (I) comprising a  
CC sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
CC having the biological activity of MCP-1. A protein encoded by a nucleic  
CC acid of the invention has cytostatic, and antiarteriosclerotic activity.  
CC A nucleic acid of the invention may have a use in gene therapy. A  
CC compound of the invention is useful in the preparation of a medicament  
CC for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

Tue Aug 9 17:01:26 2005

CC pharmaceutical composition of the invention is useful for the treatment  
CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
CC atherosclerosis or cancer. The present sequence is used in the  
CC exemplification of the invention.

XX  
XX  
SQ Sequence 250 BP; 56 A; 60 C; 58 G; 76 T; 0 U; 0 Other;

Query Match 100.0%; Score 250; DB 6; Length 250;  
Best Local Similarity 100.0%; Pred. No. 7.9e-73;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAATGGATTAAATGCAATGTCA 60  
Db 1 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAATGGATTAAATGCAATGTCA 60  
Qy 61 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGGCT 120  
Db 61 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGGCT 120  
Qy 121 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTTCATCTAGTTTCTCGCTTCTTCTCTTT 180  
Db 121 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTTCATCTAGTTTCTCGCTTCTTCTCTTT 180  
Qy 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTTCAATC 240  
Db 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTTCAATC 240  
Qy 241 TGTGGTCAGT 250  
Db 241 TGTGGTCAGT 250

RESULT 2  
ADE84876  
ID ADE84876 standard; DNA; 681 BP.  
XX  
XX ADE84876;  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Farnesyl transferase inhibitor modulated leukemia associated gene #95.  
XX ss; cytosolic; farnesyl transferase inhibitor; gene expression;  
XX quinolinone; leukemia; cancer.  
XX  
XX Homo sapiens.  
XX WO2003038129-A2.  
XX  
XX 08-MAY-2003.  
XX  
XX 30-OCT-2002; 2002WO-US034784.  
XX  
XX 30-OCT-2001; 2001US-0338997P.  
XX 30-OCT-2001; 2001US-0340081P.  
XX 30-OCT-2001; 2001US-0340938P.  
XX 30-OCT-2001; 2001US-0341012P.  
XX  
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
XX  
XX Rapani M;  
XX  
XX WPI; 2003-513497/48.  
XX  
XX Determining whether a patient will respond to treatment with a farnesyl  
XX transferase inhibitor, by analyzing the expression of gene that is  
XX differentially modulated in the presence of the inhibitor.  
XX  
XX Disclosure; SEQ ID NO 95; 346pp; English.  
XX  
XX The invention relates to a method of determining whether a patient will  
XX respond to treatment with a farnesyl transferase inhibitor (FTI), by  
XX analyzing the expression of gene that is differentially modulated in the

CC presence of an FTI. The method is useful for determining whether a  
CC patient will respond to treatment with a FTI such as (B)-6-(amino(4-  
CC chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl)-4-(3-chlorophenyl)-1-  
CC methyl-2-(1H)quinolinone, monitoring the therapy of a patient, treating a  
CC patient with leukemia with FTI if the analysis indicates that the patient  
CC will respond. This sequence corresponds to a gene whose expression may be  
CC modulated in the presence of FTI.

XX  
XX  
SQ Sequence 681 BP; 153 A; 205 C; 140 G; 183 T; 0 U; 0 Other;

Query Match 100.0%; Score 250; DB 10; Length 681;  
Best Local Similarity 100.0%; Pred. No. 1.2e-72;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAATGGATTAAATGCAATGTCA 60  
Db 40 AGGCTTCTATGATGCTACTATTCTGCATTTGAATGAGCAATGGATTAAATGCAATGTCA 99  
Qy 61 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGGCT 120  
Db 100 GGGAGCGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGGCT 159  
Qy 121 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTTCATCTAGTTTCTCGCTTCTTCTCTTT 180  
Db 160 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTTCATCTAGTTTCTCGCTTCTTCTCTTT 219  
Qy 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTTCAATC 240  
Db 220 TCTGCAGTTTTCGCTTTCACAGAAAGCAGATCCTTAAATAATACCTCTTAGTTTCAATC 279  
Qy 241 TGTGGTCAGT 250  
Db 280 TGTGGTCAGT 289

RESULT 3  
ADN12120  
ID ADN12120 standard; DNA; 3221 BP.  
XX  
XX ADN12120;  
XX  
XX 17-JUN-2004 (first entry)  
XX  
XX MCP1 gene promoter region.  
XX major histocompatibility class I; MHC-I; MHC-II; Cytostatic;  
XX EBV-associated cancer; Hodgkin's lymphoma; nasopharyngeal carcinoma;  
XX gastric carcinoma; Burkitt's lymphoma; T-cell lymphoma; B-cell lymphoma;  
XX parotid carcinoma; breast carcinoma; leiomyosarcoma; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO2004027036-A2.  
XX  
XX 01-APR-2004.  
XX  
XX 19-SEP-2003; 2003WO-US029684.  
XX  
XX 19-SEP-2002; 2002US-0411990P.  
XX  
XX (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
XX  
XX Ambinder RF, Yang Y, Borrello IM, Levitsky HI;  
XX  
XX WPI; 2004-295406/27.  
XX  
XX New human cell line modified to comprise and express genes encoding  
XX immunomodulators and an antigen of Epstein-Barr virus (EBV), useful for  
XX inducing or stimulating an immune response in a human to EBV-associated  
XX cancer.  
XX  
XX Disclosure; SEQ ID NO 13; 218pp; English.

CC The present invention relates to a human cell line, which lacks major  
 CC histocompatibility class I (MHC-I) and MHC-II antigens and which has been  
 CC modified to comprise and express a gene encoding an immunomodulator and a  
 CC gene encoding an antigen of Epstein-Barr virus (EBV). The human cell  
 CC line, compositions and methods are useful for inducing or stimulating an  
 CC immune response in a human to an EBV-associated cancer, where the human  
 CC has or is at risk for Hodgkin's lymphoma, nasopharyngeal carcinoma,  
 CC gastric carcinoma, Burkitt's lymphoma, T-cell lymphoma, B-cell lymphoma,  
 CC parotid carcinoma, breast carcinoma, and leiomyosarcoma. The present  
 CC sequence represents a nucleotide sequence associated with the cell line  
 CC of the invention.

XX Seq Sequence 3221 BP; 859 A; 793 C; 727 G; 842 T; 0 U; 0 Other;

Query Match 100.0%; Score 250; DB 12; Length 3221;

Best Local Similarity 100.0%; Pred. No. 2.5e-72;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATTGTCA 60

DB 2247 AGGCTTCTATGATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATTGTCA 2306

QY 61 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGGCCCT 120

DB 2307 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGGCCCT 2366

QY 121 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 180

DB 2367 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 2426

QY 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAAATAACCCCTTTAGTTTCACATC 240

DB 2427 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAAATAACCCCTTTAGTTTCACATC 2486

QY 241 TGTGGTCAGT 250

DB 2487 TGTGGTCAGT 2496

RESULT 4

ADO03802

ID ADO03802 standard; DNA; 3221 BP.

AC ADO03802;

XX 12-AUG-2004 (first entry)

DT 12-AUG-2004 (first entry)

XX Human Ccl2 promoter region used for gene knockout animal models SeqID 3.

DE human; ds; animal model; age-related macular degeneration; AMD;

XX gene knockout; Ccl2-deficient; Ccr2-deficient; drusen;

KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;

KW choroidal neovascularisation; ophthalmological; gene therapy; promoter.

XX Homo sapiens.

OS WO2004041160-A2.

XX 21-MAY-2004.

XX 16-OCT-2003; 2003WO-US032933.

XX 30-OCT-2002; 2002US-0422096P.

XX (KENT ) UNIV KENTUCKY RES FOUND.

PA Ambati J;

XX WPI; 2004-400512/37.

XX Testing candidate drug for treating age-related macular degeneration, by

PT administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and

PT analyzing development or regression of drusen and/or lipofuscin

PT accumulation in eye.

XX Disclosure; SEQ ID NO 3; 64pp; English.

XX This invention relates to a novel methods and animal models for testing

CC candidate drugs that can be used for the treatment or prevention of age-

CC related macular degeneration (AMD). Specifically, it refers to

CC administering a candidate drug to gene knockout mice, in particular Ccl2-

CC deficient, Ccr2-deficient and/ or a Ccl2-deficient/Ccr2-deficient dual

CC knockout mouse. The present invention describes analysing the knockout

CC mouse eye for development or regression of drusen and/ or lipofuscin

CC accumulation, as well as for the effect of a candidate drug on Bruch's

CC membrane, retinal degeneration and/ or choroidal neovascularisation.

CC Accordingly, such compositions exhibit ophthalmological activities and

CC can be used for gene therapy purposes. This polynucleotide sequence is

CC the human Ccl2 promoter region DNA of the invention.

XX Seq Sequence 3221 BP; 859 A; 793 C; 727 G; 842 T; 0 U; 0 Other;

Query Match 100.0%; Score 250; DB 12; Length 3221;

Best Local Similarity 100.0%; Pred. No. 2.5e-72;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATTGTCA 60

DB 2247 AGGCTTCTATGATGCTACTATTCTGCATTGGAATGAGCAAAATGGATTAAATGCAATTGTCA 2306

QY 61 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGGCCCT 120

DB 2307 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGGCCCT 2366

QY 121 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 180

DB 2367 GAAGGTAAGCTGGCAGCGAGCTTGACATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTT 2426

QY 181 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAAATAACCCCTTTAGTTTCACATC 240

DB 2427 TCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAAAAATAACCCCTTTAGTTTCACATC 2486

QY 241 TGTGGTCAGT 250

DB 2487 TGTGGTCAGT 2496

RESULT 5

ADO03803

ID ADO03803 standard; DNA; 11793 BP.

XX ADO03803;

XX 12-AUG-2004 (first entry)

DT 12-AUG-2004 (first entry)

XX Human Ccl2 gene and enhancer region DNA SeqID 4.

DE human; ds; animal model; age-related macular degeneration; AMD;

KW gene knockout; Ccl2-deficient; Ccr2-deficient; drusen;

KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;

KW choroidal neovascularisation; ophthalmological; gene therapy.

XX Homo sapiens.

OS WO2004041160-A2.

XX 21-MAY-2004.

XX 16-OCT-2003; 2003WO-US032933.

XX 30-OCT-2002; 2002US-0422096P.

XX (KENT ) UNIV KENTUCKY RES FOUND.

PA Ambati J;

XX WPI; 2004-400512/37.

XX Testing candidate drug for treating age-related macular degeneration, by

PT administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and

PT analyzing development or regression of drusen and/or lipofuscin

DR WPI; 2004-400512/37.

XX Testing candidate drug for treating age-related macular degeneration, by

XX administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and

PT analyzing development or regression of drusen and/or lipofuscin

PT accumulation in eye.

XX

XX Disclosure; SEQ ID NO 4; 64pp; English.

XX

XX This invention relates to a novel methods and animal models for testing

XX candidate drugs that can be used for the treatment or prevention of age-

CC related macular degeneration (AMD). Specifically, it refers to

CC administering a candidate drug to gene knockout mice, in particular Ccl2-

CC deficient, Ccr2-deficient and/ or Ccl2-deficient/Ccr2-deficient dual

CC knockout mouse. The present invention describes analysing the knockout

CC mouse eye for development or regression of drusen and/ or lipofuscin

CC accumulation, as well as for the effect of a candidate drug on Bruch's

CC membrane, retinal degeneration and/ or choroidal neovascularisation.

CC Accordingly, such compositions exhibit ophthalmological activities and

CC can be used for gene therapy purposes. This polynucleotide sequence is

CC the human Ccl2 gene and enhancer region DNA of the invention.

XX

XX Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;

XX

Query Match 100.0%; Score 250; DB 12; Length 11793;

Best Local Similarity 100.0%; Pred. No. 4.5e-72; Indels 0; Gaps 0;

Matches 250; Conservative 0; Mismatches 0;

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAATGGATTTAATGCAATGTCA 60

DB 7050 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAATGGATTTAATGCAATGTCA 7109

QY 61 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCTCGCTGGGAGGCCCTTGGAAATGGGCT 120

DB 7110 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCTCGCTGGGAGGCCCTTGGAAATGGGCT 7169

QY 121 GAAGGTAAGCTGGCAGGAGCTGCACATGCTTTTCATCTAGTTTCCGCTTCCTTCCTTT 180

DB 7170 GAAGGTAAGCTGGCAGGAGCTGCACATGCTTTTCATCTAGTTTCCGCTTCCTTCCTTT 7229

QY 181 TCTGCAGTTTGGCTTTCACAGAAAGCAGATCCTTAAAAATAACCTCTTAGTTTCACATC 240

DB 7230 TCTGCAGTTTGGCTTTCACAGAAAGCAGATCCTTAAAAATAACCTCTTAGTTTCACATC 7289

QY 241 TGTGGTCACT 250

DB 7290 TGTGGTCACT 7299

RESULT 6

ABK47979

ID ABK47979 standard; DNA; 2775 BP.

XX

XX ABK47979;

XX

XX 02-JUL-2002 (first entry)

XX Human small inducible cytokine A2 (SCYA2) genomic DNA.

XX Human; small inducible cytokine A2; SCYA2; gene; db; haplotype pair;

KW haplotyping; atherosclerosis; antiarteriosclerotic; gene therapy;

KW single nucleotide polymorphism; genotyping; drug screening;

KW chromosome 17q11.2-q21.1.

XX

XX Homo sapiens.

XX

XX Key Location/Qualifiers

PH 598..2080

FT /\*tag= a

FT /product= "Human SCYA2"

XX

XX WO200218413-A2.

XX

07-MAR-2002.

28-AUG-2001; 2001WO-US026899.

XX

28-AUG-2000; 2000US-0228496P.

PR

XX (GENA-) GENAISSANCE PHARM INC.

PA

XX Anastasio AE, Finkel K, Koshy B, Kumar AM, Lee HH;

PI

XX WPI; 2002-339655/37.

XX P-PSDB; AAU77179.

DR

XX New genetic variants having polymorphisms in the small inducible cytokine

PT A1 (SCYA2 ) gene, useful for studying the function of SCYA2 and for

PT treating disorders affected by expression or function of the SCYA2

PT isogene.

XX

XX Claim 1; Fig 1; 58pp; English.

XX The invention relates to single nucleotide polymorphisms in the gene

CC encoding human small inducible cytokine A2 (SCYA2) polypeptide. A method

CC for haplotyping the SCYA2 gene in an individual comprises identifying the

CC nucleotide at one or more polymorphic sites and determining whether one

CC of the copies of the gene is defined by one of the SCYA2 haplotypes given

CC in the specification or whether both copies are defined by a haplotype

CC pair. This method is useful in genotyping, whereby all possible haplotype

CC trait and a haplotype or haplotype pair of the SCYA2 gene can be

CC identified by comparing the frequency of the haplotype or haplotype pair

CC in a population exhibiting the trait with the frequency of the haplotype

CC or haplotype pair in a reference population, where a higher haplotype

CC frequency in the trait population indicates the trait is associated with

CC the haplotype or haplotype pair. SCYA2 and its corresponding DNA are used

CC for studying the expression and function of SCYA2, and in screening for

CC candidate drugs to treat diseases related to SCYA2 activity, such as

CC atherosclerosis. This sequence represents genomic DNA which encodes the

CC human SCYA2 polypeptide

XX

XX Sequence 2775 BP; 699 A; 723 C; 565 G; 778 T; 0 U; 10 Other;

XX

Query Match 95.0%; Score 237.4; DB 6; Length 2775;

Best Local Similarity 99.2%; Pred. No. 3.7e-68;

Matches 249; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAATGGATTTAATGCAATGTCA 60

DB 35 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAATGGATTTAATGCAATGTCA 94

QY 61 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCTCGCTGGGAGGCCCTTGGAAATGGGCT 120

DB 95 GGGAGCCGGCCAAAGCTTGAGAGCTCTTCTCGCTGGGAGGCCCTTGGAAATGGGCT 154

QY 121 GAAGGTAAGCTGGCAGGAGCCCTGCATGCTTTTCATCTAGTTTCCGCTTCCTTCCTTT 180

DB 155 GAAGGTAAGCTGGCAGGAGCCCTGCATGCTTTTCATCTAGTTTCCGCTTCCTTCCTTT 214

QY 181 T-CTGCAAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTTAAAAATAACCTCTTAGTTTCACAT 239

DB 215 TCTGCAAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTTAAAAATAACCTCTTAGTTTCACAT 274

QY 240 CTGTGGTCACT 250

DB 275 CTGTGGTCACT 285

RESULT 7

ABL33653/c

ID ABL33653 standard; DNA; 5926 BP.

XX

XX ABL33653;

XX

XX 26-MAR-2002 (first entry)

XX

XX DE Human immune system associated gene SEQ ID NO: 1626.  
XX DE  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; anti-naemic; cytosine; cytosine; cytosine;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;  
DE acute myeloid leukaemia; Alzheimer's disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
XX ds.  
XX Homo sapiens.  
XX WO200200928-A2.  
XX 03-JAN-2002.  
XX 02-JUL-2001; 2001WO-EP007537.  
XX 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX (EPIG-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX Claim 1; SEQ ID NO 1626; 32pp + Sequence Listing; German.  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
XX Sequence 5926 BP; 1554 A; 55 C; 1492 G; 2825 T; 0 U; 0 Other;  
Query Match 64.9%; Score 162.2; DB 6; Length 5926;  
Best Local Similarity 78.5%; Pred. No. 6.1e-43;  
Matches 194; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 4 CTTCTATGATGCTTACTTCTGATTTGAATGAGCAATGGATTAAATGCTATGTCAGG 63  
Db 1430 CTTCTATGATGCTTACTTCTGATTTGAATGAGCAATGGATTAAATGCTATGTCAGG 1371  
QY 64 AGCGGCGCAAGCTTCGAGAGCTCTTCTGCTGGGAGGCGCCCTTGGAAATGGGCTGAA 123  
Db 1370 AACCGACCAAACTTAAABACTCTCTTCTAATTAATAAACCCTTAAATATAACCTAA 1311  
QY 124 GGTAAGCTGGGAGCGAGCTGATGCTTTCTAGTTCTCTGCTTCTCTCTCTCTCTCTCT 183  
Db 1310 AATAAACTAACACGACCACTTCTATCTAATTTCTGCTGCTCTCTCTCTCTCTCTCT 1251  
QY 184 GCAGTTTTCGCTTTCACAGAAAGCAGAACTCTTAAATAACCTCTCTAGTTTCACATCTGT 243  
Db 1250 ACAATTTTCGCTTTCACAGAAAGCAGAACTCTTAAATAACCTCTCTAGTTTCACATCTGT 1191  
QY 244 GGTCACT 250  
Db 1190 AATCAAT 1184

RESULT 8  
ABL33652

ID ABL33652 standard; DNA; 5926 BP.  
XX ABL33652;  
XX 26-MAR-2002 (first entry)  
XX Human immune system associated gene SEQ ID NO: 1625.  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; anti-naemic; cytosine; cytosine; cytosine;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;  
KW acute myeloid leukaemia; Alzheimer's disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
XX ds.  
XX Homo sapiens.  
XX WO200200928-A2.  
XX 03-JAN-2002.  
XX 02-JUL-2001; 2001WO-EP007537.  
XX 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX (EPIG-) EPIGENOMICS AG.  
XX Olek A, Piepenbrock C, Berlin K;  
XX WPI; 2002-130909/17.  
XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX Claim 1; SEQ ID NO 1625; 32pp + Sequence Listing; German.  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
XX Sequence 5926 BP; 1562 A; 55 C; 1318 G; 2991 T; 0 U; 0 Other;  
Query Match 64.2%; Score 160.4; DB 6; Length 5926;  
Best Local Similarity 77.6%; Pred. No. 2.4e-42;  
Matches 194; Conservative 0; Mismatches 56; Indels 0; Gaps 0;  
QY 1 AGGCTTCTATGATGCTTACTTCTGATTTGAATGAGCAATGGATTAAATGCTATGTCAG 60  
Db 4494 AGGCTTCTATGATGCTTACTTCTGATTTGAATGAGCAATGGATTAAATGCTATGTCAG 4553  
QY 61 GGGAGCGCGCAAGCTTCGAGAGCTCTTCTGCTGGGAGGCGCCCTTGGAAATGGGCT 120  
Db 4554 GGGAGCTCGTTAAAGTTTGAGAGTTTTTTTTTTGGTTGGAGGTTTTTGGAAATGGGTTT 4613  
QY 121 GAGGTAAGCTGGGAGCGAGCTGATGCTTTTCTAGTTTCTCTGCTCTCTCTCTCTCTCT 180  
Db 4614 GAGGTAAGCTGGGAGCGAGCTGATGCTTTTCTAGTTTCTCTGCTCTCTCTCTCTCTCT 4673  
QY 181 TCTGAGTTTTCGCTTTCACAGAAAGCAGAACTCTTAAATAACCTCTCTAGTTTCACATC 240  
Db 4674 TTTGAGTTTTCGCTTTCACAGAAAGCAGAACTCTTAAATAACCTCTCTAGTTTCACATC 4733  
QY 241 TGTGTCAGT 250

07-OCT-2004 (first entry)  
 Mouse cancer associated sequence MD08-030, SEQ ID 287.  
 Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.  
 Mus musculus.  
 WO2004060304-A2.  
 22-JUL-2004.  
 22-DEC-2003; 2003WO-US041389.  
 27-DEC-2002; 2002US-00330773.  
 (SAGR-) SAGRES DISCOVERY INC.  
 Morris DW, Malandro MS;  
 WPI; 2004-543781/52.  
 New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.  
 Claim 1; SEQ ID NO 287; 199pp; English.  
 The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic formate directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 Query Match 14.6%; Score 36.4; DB 12; Length 80275;  
 Best Local Similarity 51.2%; Pred. No. 1.7; Indels 0; Gaps 0;  
 Matches 85; Conservative 0; Mismatches 81;  
 60 AGGAGCGCGGCGAAAGCTTCAGAGCTTCCTTCCTGCTGGAGGCCCTTGAATGTGGCC 119  
 67002 AAGAGCTTCCTTAGCTTGGGCTCATCTGAGTTCGATGGGATGGGTGGCAGGCC 119  
 120 TGAAGTGAAGTGGCAGCAGCTTCATCTTCATCTAGTTTCCTCGCTTCCTTCCTT 179  
 66942 TGGCTGAAAACTCCAGCCCTTCTCACCGATGTTCTGCAAGACATCATTCTTGT 66883  
 180 TTCTCAGTTTTCGCTTCACAGAACAGAAATCCTTAAATAAAC 225  
 66882 CTTCGCTTATCATCTTAGCCACTCAATGGCTCATAAATAAAC 66837

RESULT 11  
 ABD33157/c  
 ID ABD33157 standard; DNA; 109559 BP.  
 XX  
 AC ABD33157;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Murine cancer-associated (CA) gene MD07-021.  
 XX  
 KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;  
 KW ds; cancer; cytostatic.  
 XX  
 OS Mus musculus.  
 XX  
 WO2004058146-A2.  
 XX  
 15-JUL-2004.  
 XX

07-OCT-2004 (first entry)  
 Mouse cancer associated sequence MD08-030, SEQ ID 287.  
 Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.  
 Mus musculus.  
 WO2004060304-A2.  
 22-JUL-2004.  
 22-DEC-2003; 2003WO-US041389.  
 27-DEC-2002; 2002US-00330773.  
 (SAGR-) SAGRES DISCOVERY INC.  
 Morris DW, Malandro MS;  
 WPI; 2004-543781/52.  
 New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.  
 Claim 1; SEQ ID NO 287; 199pp; English.  
 The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic formate directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 Query Match 14.6%; Score 36.4; DB 12; Length 80275;  
 Best Local Similarity 51.2%; Pred. No. 1.7; Indels 0; Gaps 0;  
 Matches 85; Conservative 0; Mismatches 81;  
 60 AGGAGCGCGGCGAAAGCTTCAGAGCTTCCTTCCTGCTGGAGGCCCTTGAATGTGGCC 119  
 67002 AAGAGCTTCCTTAGCTTGGGCTCATCTGAGTTCGATGGGATGGGTGGCAGGCC 119  
 120 TGAAGTGAAGTGGCAGCAGCTTCATCTTCATCTAGTTTCCTCGCTTCCTTCCTT 179  
 66942 TGGCTGAAAACTCCAGCCCTTCTCACCGATGTTCTGCAAGACATCATTCTTGT 66883  
 180 TTCTCAGTTTTCGCTTCACAGAACAGAAATCCTTAAATAAAC 225  
 66882 CTTCGCTTATCATCTTAGCCACTCAATGGCTCATAAATAAAC 66837

RESULT 11  
 ABD33157/c  
 ID ABD33157 standard; DNA; 109559 BP.  
 XX  
 AC ABD33157;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Murine cancer-associated (CA) gene MD07-021.  
 XX  
 KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;  
 KW ds; cancer; cytostatic.  
 XX  
 OS Mus musculus.  
 XX  
 WO2004058146-A2.  
 XX  
 15-JUL-2004.  
 XX

4734 TGTGTTAGT 4743  
 RESULT 9  
 ACF64398 standard; DNA; 800 BP.  
 XX  
 AC ACF64398;  
 XX  
 DT 13-OCT-2003 (first entry)  
 XX  
 DE Human MCP1 nucleotide sequence >MCP1\_01.  
 XX  
 KW Human; detection; computer-readable storage medium; polymorphic site;  
 KW signal carrying data; data processing system; multiple sclerosis; gene;  
 KW ds.  
 XX  
 OS Homo sapiens.  
 XX  
 WO2003014319-A2.  
 XX  
 20-FEB-2003.  
 XX  
 07-AUG-2002; 2002WO-US025268.  
 XX  
 07-AUG-2001; 2001US-0310741P.  
 PR  
 24-SEP-2001; 2001US-0324790P.  
 XX  
 (DNAS-) DNA SCI INC.  
 PA  
 Jones HB, Xu H, White R, Rienhoff HY, Jin W, Natsoulis G;  
 FI  
 WPI; 2003-268196/26.  
 XX  
 New polynucleotide, useful for detecting loci associated with multiple sclerosis.  
 Disclosure; Page 74; 93pp; English.  
 XX  
 The present invention describes an isolated polynucleotide (PN) comprising: (a) a sequence comprising at least 15 contiguous nucleotides of a sequence comprising variant sequences (A) from Table 4 given in the specification; or (b) a sequence that is complementary to (A). Also described: (1) an array of (PN)s comprising two or more of the isolated (PN)s; (2) detecting a (PN) in an individual; (3) a computer-readable storage medium, where each record has a field identifying a base occupying a (PN) site and a location of the polymorphic site; and (4) a signal carrying data for access by an application program having executed on a data processing system. The (PN) can be used for detecting loci associated with multiple sclerosis. ACF64025 to ACF64424 represent sequences used in the exemplification of the present invention  
 Query Match 36.6%; Score 91.4; DB 8; Length 800;  
 Best Local Similarity 98.9%; Pred. No. 9.8e-20;  
 Matches 92; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 158 TAGTTTCCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTAA 217  
 1 TAGTTTCCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTAA 60  
 218 AAATAACCTCTTAGTTTCACATCTGTGTCAGT 250  
 61 AAATAACCTCTTAGTTTCACATCTGTGTCAGT 93

RESULT 10  
 ADQ97310/c  
 ID ADQ97310 standard; DNA; 80275 BP.  
 XX  
 AC ADQ97310;  
 XX

```
PF 15-DEC-2003; 2003WO-US040081.
XX
PR 17-DEC-2002; 2002US-00322281.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-499109/47.
XX
PT Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.
XX
PS Disclosure; SEQ ID NO 137; 182pp; English.
XX
CC The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published\_pct\_sequences
XX
SQ Sequence 109559 BP; 31165 A; 24061 C; 22798 G; 30464 T; 0 U; 1071 Other;

Query Match 14.6%; Score 36.4; DB 13; Length 109559;
Best Local Similarity 55.6%; Pred. No. 2;
Matches 90; Conservative 0; Mismatches 71; Indels 1; Gaps 1;

QY 36 AGCAATGGATTTAATGCATTGTGAGGAGCGCGCCAAAGCTTGAGAGCTCTCTCTGGC 95
DB 106445 AGGAGCTTGCTCGCATGGTTGTGAAGGAAGGCCAGGTGTGCCAGGCTGTCTCGAGGC 106386
QY 96 TGGGAGGCGCCCTTGGGAATGTGGCCTGAAGTGAAGCTGGCAGGAGCCTGACATGCTTTCA 155
DB 106385 TGGGAGGAGCCTAGGCAT-TGGGATGGAGGAGCAGGGTTGTAGGCTTACCTAGCTCCATTCT 106327
QY 156 TCTAGTTTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 197
DB 106326 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 106285

RESULT 12
ABL86846/c
ID ABL86846 standard; cDNA; 165 BP.
XX
AC ABL86846;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:9824.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX

PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
XX
Claim 1; SEQ ID NO 9824; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
SQ Sequence 165 BP; 67 A; 37 C; 38 G; 22 T; 0 U; 1 Other;

Query Match 14.3%; Score 35.8; DB 6; Length 165;
Best Local Similarity 53.7%; Pred. No. 0.16;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 87 CTTCTGCTGGGAGGCCCTTGGGAATGTGGCCTGAAGGTAACTGGCAGCGAGCTGAC 146
DB 145 CTCCTGCGAGGTGTGCCCCCTTGGCCTGGACCTGGGGCCTGAATTGTGGGAAGGTGGT 86
QY 147 ATGCTTTTCATCTAGTTTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 206
DB 85 TCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26
QY 207 AGAATCCTTAAATAAATA 222
DB 25 AGAATTAATACNCTA 10

RESULT 13
AAS88382
ID AAS88382 standard; cDNA; 1500 BP.
XX
AC AAS88382;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #24186.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PR 30-MAR-2001; 2001WO-US008631.
```

Tue Aug 9 17:01:26 2005

us-09-899-276c-4.rng

```
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG24195.
DR New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 24186; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have application of mutations
CC diagnostic, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1500 BP; 292 A; 380 C; 440 G; 388 T; 0 U; 0 Other;
SQ
Query Match 13.6%; Score 34; DB 5; Length 1500;
Best Local Similarity 50.6%; Pred. No. 1.8;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 37 GCAATGGATTAAATGCATTGTTCAGGAGCCGCGCAAGCTTGAGAGCTCTCTCGCT 96
Db 1302 GCACGTACATATCAGCAGCGCCAGCCGCTCCGCTGGAAGCTGTGGAGTACATTTCTCGCA 1361
QY 97 GGGAGGCCCTTGGAAATGTGGCTGAAGTAAGCTGGAGCGAGCGAGCTGACATGCTTTCAT 156
Db 1362 CCTTGCCAGCTCGAAAAAGGGAGAGGGTCTGCGGGCAAGTCGACCGCGCGCTACTA 1421
QY 157 CTAGTTTCCTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 198
Db 1422 ATAAAGCATCAGGATTCCTGCTATCGTGAGCTTTCCTCTCTCTCTCTCTCTCTCTCTCA 1463
RESULT 14
AAK73223/c
ID AAK73223 standard; DNA; 31051 BP.
XX AC
XX AAK73223;
XX AC
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28035.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
```

```
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225466P.
PR 14-AUG-2000; 2000US-0225467P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
```

27-SEP-2000; 2000US-0235836P.  
29-SEP-2000; 2000US-0236327P.  
29-SEP-2000; 2000US-0236367P.  
29-SEP-2000; 2000US-0236368P.  
29-SEP-2000; 2000US-0236369P.  
29-SEP-2000; 2000US-0236370P.  
02-OCT-2000; 2000US-0236802P.  
02-OCT-2000; 2000US-0237037P.  
02-OCT-2000; 2000US-0237038P.  
02-OCT-2000; 2000US-0237039P.  
02-OCT-2000; 2000US-0237040P.  
13-OCT-2000; 2000US-0239935P.  
13-OCT-2000; 2000US-0239937P.  
20-OCT-2000; 2000US-0239937P.  
20-OCT-2000; 2000US-0240960P.  
20-OCT-2000; 2000US-0241221P.  
20-OCT-2000; 2000US-0241785P.  
20-OCT-2000; 2000US-0241786P.  
20-OCT-2000; 2000US-0241787P.  
20-OCT-2000; 2000US-0241808P.  
20-OCT-2000; 2000US-0241809P.  
20-OCT-2000; 2000US-0241826P.  
01-NOV-2000; 2000US-0244617P.  
08-NOV-2000; 2000US-0246474P.  
08-NOV-2000; 2000US-0246475P.  
08-NOV-2000; 2000US-0246476P.  
08-NOV-2000; 2000US-0246477P.  
08-NOV-2000; 2000US-0246478P.  
08-NOV-2000; 2000US-0246523P.  
08-NOV-2000; 2000US-0246524P.  
08-NOV-2000; 2000US-0246525P.  
08-NOV-2000; 2000US-0246526P.  
08-NOV-2000; 2000US-0246527P.  
08-NOV-2000; 2000US-0246528P.  
08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246609P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249219P.  
17-NOV-2000; 2000US-0249245P.  
17-NOV-2000; 2000US-0249246P.  
17-NOV-2000; 2000US-0249265P.  
17-NOV-2000; 2000US-0249297P.  
17-NOV-2000; 2000US-0249299P.  
17-NOV-2000; 2000US-0249300P.  
01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
05-DEC-2000; 2000US-0251030P.  
05-DEC-2000; 2000US-0251988P.  
05-DEC-2000; 2000US-0251989P.  
05-DEC-2000; 2000US-0251990P.  
11-DEC-2000; 2000US-0254097P.  
05-JAN-2001; 2001US-0259678P.  
(HUMA-) HUMAN GENOME SCI INC.  
XX  
XX

PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
DR  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Disclosure; SEQ ID NO 28035; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins, and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent  
CC diagnosis and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX Sequence 31051 BP; 6999 A; 8773 C; 8261 G; 7018 T; 0 U; 0 Other;  
SQ

Query Match 13.5%; Score 33.8; DB 4; Length 31051;  
Best Local Similarity 56.9%; Pred. No. 8.2;  
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
QY 30 TGAATGAGCAATGGATTATGCAATTGTCAGGAGCCGCGCAAGCTTGAGAGCTCCTT 89  
Db 27616 TGACTAGGGAATTCCTTAACCTGCTTATCTGCAATGGCCCAATTTAAGGTTCTCTT 27557  
QY 90 CCTGGCTGGAGGCCCTTGGAAATGCTGGAATGCTGAAGTAAAGCTGGCAGCG 138  
Db 27556 CCAGCCCTCAAAACTCTATGATCAGAGGCACTTGGTGGTCTCTGATG 27508

RESULT 15  
ABK35434  
ID ABK35434 standard; cDNA; 673 BP.  
XX  
AC ABK35434;  
XX  
DT 08-MAY-2002 (first entry)  
XX Human cDNA encoding secreted protein #572.  
XX  
DE Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;  
KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
KW tissue regeneration; wound healing; burn; haematopoiesis;  
KW myeloid cell deficiency; lymphoid cell deficiency.  
XX Homo sapiens.  
XX  
XX WO200177288-A2.  
XX  
XX 18-OCT-2001.  
XX  
XX 29-MAR-2001; 2001WO-US010224.  
XX  
XX 06-APR-2000; 2000US-0195582P.  
XX (GEMY ) GENETICS INST INC.  
XX



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 1179.96 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-4  
Perfect score: 250  
Sequence: 1 aggtctctatgatgctacta.....agttcacatctgtgtcagt 250

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	68.8	707	8	BZ598459 WHABX44TR
2	39.8	15.9	768	4	BG121617 602351571
3	37	14.8	437	1	AA568299 nf15405.8
4	36.8	14.7	161	1	AI962277 wq46d11.x
5	36	14.4	371	8	AZ997556 2M0284F07
6	36	14.4	481	8	AZ536871 Gm I8B001
7	35.8	14.3	165	1	AI280614 qu03h06.x
8	35.6	14.2	391	9	AG264604 Lotus cor
9	35.6	14.2	571	9	CL579550 OB_Ra003
10	35.2	14.1	200	1	AI950785 wx55b02.x
11	35.2	14.1	378	1	AI208450 qg36c09.x
12	35.2	14.1	396	1	AI302643 qn48a06.x
13	35.2	14.1	683	5	BU616089 UI-H-DF0
14	35	14.0	641	9	CR841822 GR0AAA76A
15	35	14.0	661	8	BH325278 CH230-45J
16	35	14.0	676	8	BZ257915 CH230-479
17	34.6	13.8	606	4	B1321217 saf49a08.
18	34.4	13.8	554	6	CD478218 eca01-53m
19	34.4	13.8	576	6	CD306129 StrPu691.
20	34.4	13.8	811	8	BH603026 B0GJU24TF
21	34.2	13.7	416	8	BZ179889 CH230-444
22	34.2	13.7	540	8	BH276950 CH230-48M
23	33.8	13.5	171	6	CD349976 UI-M-FY0-
24	33.8	13.5	633	2	BB612111 BB612111

C	25	33.8	13.5	634	6	BY727711	BY727711
	26	33.8	13.5	636	9	CE785192	tigr-g88-
	27	33.8	13.5	638	6	BY727766	BY727766
	28	33.8	13.5	645	2	BB645191	BB645191
	29	33.8	13.5	671	2	BB646050	BB646050
	30	33.8	13.5	749	8	BH267859	CH230-66L
	31	33.8	13.5	1131	5	BQ948968	AGENCOURT
	32	33.6	13.4	872	4	BJ747052	BJ747052
	33	33.4	13.4	418	1	AI639337	rx010168
	34	33.4	13.4	477	1	AI096606	QB44h08.x
	35	33.4	13.4	528	8	BZ907804	CH240_31F
	36	33.4	13.4	593	9	CE635729	CE635729
	37	33.4	13.4	635	8	BH050645	tigr-g88-
	38	33.4	13.4	720	5	BU340316	RFCI-24-2
	39	33.4	13.4	772	8	AO395289	603518092
	40	33.2	13.3	488	9	CE822654	CITBI-E1-
	41	33.2	13.3	637	9	CE591509	tigr-g88-
	42	33.2	13.3	703	7	CO211547	tigr-g88-
	43	33.2	13.3	849	7	CO399237	WS00921.B
	44	33	13.2	261	8	AQ113303	AGENCOURT
	45	33	13.2	498	9	CL890773	CIT-HSP-2
							abg05d04.

ALIGNMENTS

RESULT 1  
BZ598459 WHABX44TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo sapiens genomic clone MCF7\_1-12H16, genomic survey sequence.  
LOCUS BZ598459.1 GI:31506921  
DEFINITION BZ598459  
ACCESSION BZ598459  
VERSION BZ598459.1  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 707)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
VOLIK.S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J.W. and Collins, C.  
End-sequence profiling: Sequence-based analysis of aberrant genomes  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
MEDLINE 22709111  
PUBMED 12788976  
COMMENT Contact: Volik SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
http://www.genomex.com  
Class: BAC ends.

FEATURES  
Location/Qualifiers  
1..707  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-12H16"  
/sex="female"  
/clone\_lib="Human MCF7 breast cancer cell line library (MCF7\_1)"  
/note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 68.8%; Score 172; DB 8; Length 707;  
Best Local Similarity 100.0%; Pred. No. 8.9e-42;

Tue Aug 9 17:01:26 2005

```

Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTGGCTGAAGTAAGCTGGCAGCG 138
    |||||
Db 1 GAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTGGCTGAAGTAAGCTGGCAGCG 60
    |||||

QY 139 AGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTCTTCTTCTTCTGAGTTTCGCTTCA 198
    |||||
Db 61 AGCTGACATGCTTTTCATCTAGTTTCCTCGCTTCCTTCTTCTTCTTCTGAGTTTCGCTTCA 120
    |||||

QY 199 CAGAAAGCAGATCCTTAAATATACCTCTTGTAGTTTACATCTGTGTGTCAGT 250
    |||||
Db 121 CAGAAAGCAGATCCTTAAATATACCTCTTGTAGTTTACATCTGTGTGTCAGT 172
    |||||

RESULT 2
BG121617 768 bp mRNA linear EST 30-JAN-2001
LOCUS 602351571F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4449977 5',
DEFINITION mRNA sequence.
VERSION BG121617
KEYWORDS BG121617.1 GI:12615126
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM10234 row: m column: 18
High quality sequence stop: 690.
Location/Qualifiers
1. .768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4449977"
/tissue_type="adenocarcinoma, cell line"
/lab_host="NIH MGC_90"
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 15.9%; Score 39.8; DB 4; Length 768;
Best Local Similarity 57.7%; Pred. No. 0.55; Indels 0; Gaps 0;
Matches 71; Conservative 0; Mismatches 52;

QY 23 CTGCATTGATGAGCAATGATTTTAATGATGTCAGGAGCCGCCGCAAGCTTGAGA 82
    |||||
Db 270 CTGCATTGCGTTTGCAGATGAATTCACAGATTAGCAGAGACTTCTTACATAGAGAT 329
    |||||

QY 83 GCTCTCTTCTGGCTGGAGGCCCTTGGAAATGTGGCTTGAAGTAGCTGGCAGCGAGCC 142
    |||||
Db 330 GCACATCCAGTCTCTGGTGGTTCCTCTGTGGGGCTTGATCGTGTGCTGCTGTTTACA 389
    |||||

QY 143 TGA 145
    |||
Db 390 TGA 392
    |||

```

```

RESULT 3
AA568299 437 bp mRNA linear EST 22-AUG-1997
LOCUS n115a05.s1 NCI_CGAP_Pri Homo sapiens cDNA clone IMAGE:913808, mRNA
DEFINITION sequence.
ACCESSION AA568299
VERSION AA568299.1 GI:2341353
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 394.
Location/Qualifiers
1. .437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:913808"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pri"
/notes="Vector: PAMPI0; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected, histologically normal
prostate epithelial cells. Double-stranded cDNA was
ligated to EcoRI adaptors. 5 cycles of PCR applied to the
cDNA with an adaptor-specific primer, and the resulting
PCR product subcloned into PAMPI0 by the UDG-cloning
method (Life Technologies). Average insert size is 600
bp. NOTE: Not directionally cloned. This library was
constructed by David Krizman."

ORIGIN
Query Match 14.8%; Score 37; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.4; Indels 0; Gaps 0;
Matches 37; Conservative 0; Mismatches 0;

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAG 37
    |||||
Db 401 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAG 437
    |||||

RESULT 4
AI962277/c 161 bp mRNA linear EST 20-AUG-1999
LOCUS WQ46d11.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2474325 3',
DEFINITION mRNA sequence.
ACCESSION AI962277
VERSION AI962277.1 GI:5754990
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 161)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40UP from Gibco.

## ORIGIN

[illegible]

RESULT 5	A2997556	371 bp	DNA	linear	GSS 27-APR-2001
LOCUS	A2997556				
DEFINITION	Mouse 10kb plasmid UGUC2M library Mus musculus genomic clone UGUC2M0284F07 F, genomic survey sequence.				
ACCESSION	A2997556				
VERSION	A2997556.1	GI:13868793			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 371)				
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Stokes,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.				

TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLUC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0284 row: F column: 07  
Seq primer: CGTGTGAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 371.

FEATURES  
    source Location/Qualifiers  
        1..371  
            /organism="Mus musculus"  
            /mol\_type="genomic DNA"  
            /strain="C57BL/6J"  
            /db\_xref="taxon:10090"  
            /clone="UUGC2M0284F07"  
            /sex="Female"  
            /lab\_host="E. coli strain XL10-Gold, Tl-resistant, P-"  
            /clone\_lib="Mouse 10kb plasmid UUGC2M library"  
            /note="Vector; PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (female) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 [gi|4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

ORIGIN  
  
Query Match 14.4%; Score 36; DB 8; Length 371;  
Best Local Similarity 55.6%; Pred. No. 6.6;  
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 9 ATGCATCTACTATTTCGCAATTTGAATGAGCAAATGGATTAATGCATTGTCCAGGGAGCGG 68  
|||  
Db 60 ATTCTGCTAATAATTTTCCTTGACACAGAANAATGGGTACAAATAATGTATGAGGGTGCAG 119  
|||  
QY 69 GCCAAGCTTGAGAGCTCCTTCCTGGCTGGGAGGCCCTTGGAAATGTGGCCTGAAGGTAA 128  
|||  
Db 120 GCAGAGATGAGGARAGCCTTGGTTGGTTTAGGCATCTAAAAAATCCCCAGAGAAATA 179  
|||  
QY 129 GCTG 132  
|||  
Db 180 GATG 183  
|||

RESULT 6  
AZ536871/c LOCUS  
DEFINITION  
Gm\_isb001\_082\_j15\_R ISU Soybean BAC library (pBelosAC11 HindIII)  
Glycine max genomic clone Gm\_isb001\_082\_j15, genomic survey  
sequence.  
ACCESSION AZ536871  
VERSION AZ536871.1 GI:11113450  
KEYWORDS GSS.

Tue Aug 9 17:01:26 2005

```

SOURCE      Glycine max (soybean)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1 (bases 1 to 481)
AUTHORS    Marek,L.F., Paz,M., Darnielle,L., Hanson,N. and Shoemaker,R.C.
TITLE      BAC End sequences from a soybean genomic library (ISU)
JOURNAL    Unpublished (2000)
COMMENT    Contact: Shoemaker Randy C
            Agronomy Department
            Iowa State University
            Ames, IA 50011-1010, USA
            Tel: 515 294 1205
            Fax: 515 294 2299
            Email: rcshoemaker@iastate.edu
            This BAC identified by SSR Satt129. For more information, see
            SoyBase at:
            http://genome.cornell.edu/cgi-bin/WebAce/webace?db=soybase.
            Please see as an authority for the mapping/naming: Cregan P.B., T.
            Jarvik, A.L., Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya,
            T.T. VanRoai, D.G. Lohnes, J. Chung, and J.B. Specht. 1999a. An
            integrated genetic linkage map of the soybean genome. Crop Sci.
            39: (In press)
            Seq primer: BACR or M13R
            Class: BAC ends.

FEATURES    Location/Qualifiers
            source
            1..481
            /organism="Glycine max"
            /mol_type="genomic DNA"
            /cultivar="Williams82"
            /db_xref="taxon:3847"
            /clone="Gm ISB001 082 J15"
            /tissue_type="primary_leaves"
            /dev_stage="cotyledon stage"
            /clone_lib="ISU Soybean BAC Library (pBelobAC11 HindIII)"
            /note="vector: pBelobAC11; The ISU BAC library (Marek and
            Shoemaker, Genome 40:420, 1997) was constructed using the
            HindIII site in pBelobAC11. The library consists of
            approximately 40,000 clones with an average insert size of
            150 Kb, equal to 4 haploid genome equivalents. Screening
            of the library is done by PCR amplification of DNA pools.
            "

ORIGIN
Query Match      14.4%; Score 36; DB 8; Length 481;
Best Local Similarity 55.6%; Pred. No. 7.2;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 107 TTGGAATGGCGCTGAAGTGAAGTGGGAGGAGCGCTGACATGCTTTTCATCTAGTTTCCT 166
    |||||
Db 161 TTGGTGGGAACTTGACTGTATAGAATGGCAATCACAACATGGGTTCCTCCCTCATCCT 102
    |||||

QY 167 CGCTTCCTCTCTTCTCGAGTTTTCGCTTCACAGAAAGCAGAATCCTTAAAAATAACCC 226
    |||||
Db 101 CACCCTCTTCATTTGCCCCAGTTTCTCATTCGTCACAGCAGGATGATTAATTTGCCTC 42
    |||||

QY 227 TCTT 230
    |||
Db 41 TTTT 38
    |||

RESULT 7
AI280614/c
LOCUS      AI280614.1 165 bp mRNA linear EST 21-DEC-1998
DEFINITION qu03h06.x1 NCI CGAP_Ov23 Homo sapiens cDNA clone IMAGE:1963739 3'
            similar to contains PTR5.b3 PTR5 repetitive element ;, mRNA
            sequence.
ACCESSION  AI280614
VERSION     AI280614.1 GI:3918847
KEYWORDS    EST.
SOURCE      Homo sapiens (human)

ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 165)
AUTHORS    NCI-CGAP cancer center://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapsb@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 2791 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 127.
            Location/Qualifiers
            1..165
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1963739"
            /tissue_type="tumor, 5 pooled (see description)"
            /lab_host="DH10B"
            /clone_lib="NCI-CGAP_Ov23"
            /note="Organ: ovary; Vector: pCMV-SPORT6; Site 1: SalI;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.35 kb. Tumor types include: mixed
            Mullerian tumor, papillary serous, clear cell, spindle
            cell. All are primary tumors, metastasis positive. Life
            Technologies catalog #: 11534-013"

ORIGIN
Query Match      14.3%; Score 35.8; DB 1; Length 165;
Best Local Similarity 53.7%; Pred. No. 6;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 87 CTTCTGCTGGGAGGCCCTTGGAACTGGCCCTGAAGTGAAGCTGGCAGCGAGCCTGAC 146
    |||||
Db 145 CTGCCTGCAGGTGTGGCCCTTGGCCCTGGACCTGGGCGCTGAATTGTGGAGGGTGGT 86
    |||||

QY 147 ATGCTTTTCATCTAGTTTCTCTGCTTCCTTCCTTTTCGAGTTTCGCTTCACAGAAAGC 206
    |||||
Db 85 TTCTTCTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 26
    |||||

QY 207 AGAATCCTTAAAAATA 222
    |||||
Db 25 AGAATTAATAACNCTA 10
    |||||

RESULT 8
AG264604/c
LOCUS      AG264604.1 391 bp DNA linear GSS 22-JUL-2003
DEFINITION Lotus corniculatus var. japonicus DNA, clone:LJ760008_not, genomic
            survey sequence.
ACCESSION  AG264604
VERSION     AG264604.1 GI:26664443
KEYWORDS    GSS.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM    Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
            Lotus.
REFERENCE   1
AUTHORS    Sato,S., Nakamura,Y. and Tabata,S.
TITLE      Lotus japonicus TAC End sequences
JOURNAL    Published Only in Database (2002)

```



[illegible]

RESULT 13  
 BU616089/c  
 LOCUS  
 DEFINITION  
 EST 23-SEP-2002  
 linear  
 683 bp mRNA  
 NC1 CGAP DF0 Homo sapiens CDNA clone  
 UI-H-DF0-bex-i-14-0-UI.s1  
 mRNA sequence.  
 UI-H-DF0-bex-i-14-0-UI 3',  
 ACCESSION  
 BU616089  
 VERSION  
 BU616089.1  
 KEYWORDS  
 EST.  
 SOURCE  
 GI:23282297  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 683)  
 NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 UNPUBLISHED (1997)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 The following repetitive elements were found in this cDNA  
 sequence: 63-107, >POLY\_A#Simple\_repeat  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

Tue Aug 9 17:01:26 2005

LOCUS BH325278 661 bp DNA linear GSS 03-DEC-2001  
 DEFINITION CH230-45J13.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 CH230-45J13, genomic survey sequence.

ACCESSION BH325278  
 VERSION BH325278.1 GI:17255992

KEYWORDS  
 SOURCE  
 ORGANISM

Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 661)  
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,  
 Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,  
 Riggs,F., de Jong,P. and Fraser,C.M.  
 TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: CH230-45J13.TV

Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/or ering\_information.html). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 45 row: J column: 13  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source Location/Qualifiers  
 1..661  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SSNHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-45J13"  
 /sex="female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 1"  
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
 Pieter de Jong"

ORIGIN  
 Query Match 14.0%; Score 35; DB 8; Length 661;  
 Best Local Similarity 52.4%; Pred. No. 16; Indels 0; Gaps 0;  
 Matches 77; Conservative 0; Mismatches 70;  
 QY 51 TGCATTGTGAGGAGCGCCAAAGCTTGAGAGCTCTTCCTGGTGGAGGCCCTTGG 110  
 DB 441 TGATGTGACATGATCCTGGCAAGGTAAAGTTCAAAGACCTGCTGGCAGTAGACTTGC 500  
 QY 111 AATGTGGCCTGAAGGTAAGCTGCACGAGCGCTGACATGCTTTCATCTAGTTTCCTCGCT 170  
 DB 501 AATAATGGTATTAGTACCTGCTGGTGGCTTGTAGTTGCTTGTGCTTGTGCT 560  
 QY 171 TCCTTCCTTTTCTGACGTTTTCGCTTC 197  
 DB 561 TGCTTGTGCTTGTGCTTTCCTTGC 587

Search completed: August 4, 2005, 14:32:21  
 Job time : 1187.96 secs



```

RESULT 4
US-09-949-016-58722
; Sequence 58722, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 58722
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58722

```

	Query Match Best Local Similarity Matches	12.6%; 52.8%; 65;	Score 31.4; Pred. No. 0.68; Conservative 1;	DB 4; Indels 57; Mismatched 0;	Length 601;
QY	44	GATTTATGCA	TTCTCAGGAGCGCGCCAAAGCTTGAGAGCTCTTCCTGGCTGGGAGGC	103	
DB	245	GAACACCTTCATCGTGTGTGTCAGCCAGCCTCACCTTGACAGGTCATCTCTGGCTGGC	RTAA	304	
QY	104	CCCTTGGAAATGTGGCCCTGAAGGTAAGCTGGCAGGAGCCGTGACATGCTTTCATCTAGTTT	163		
DB	305	CCCAAGGCCCAATGTGAAAGGACATTTTCATCTAGGACATGATGGGGTGCCTTGT	TTT	364	
QY	164	CCT	166		
DB	365	CTT	367		

```

RESULT 5
US-09-949-016-13807
; US-09-949-016-13807, Application US/09949016
; Sequence 13807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13807
; LENGTH: 168971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13807

```

Query Match	12.5%;	Score 31.2;	DB 4;	Length 168971;
Best Local Similarity	47.4%;	Pred. No. 24;		

```

; APPLICANT: KIM, Young-Sung
; TITLE OF INVENTION: A NOVEL CHITIN-BINDING RECEPTOR KINASE AND THE GENE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: seq
; CURRENT APPLICATION NUMBER: US/09/503,922
; CURRENT FILING DATE: 2000-02-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2943
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; US-09-503-922-2

Query Match      13.0%; Score 32.4; DB 3; Length 2943;
Best Local Similarity 54.1%; Pred. No. 0.77; 56; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 0;

QY      88  TTCCTGGCTGGGAGGCCCTTTGGAATGTGGCCCTGAAGGTAAGCTGGCAGCGAGCCTGACA 147
Db      2162  TTCCTGGAGTATAGACCTCGCTTAACGTACTCTGGAGGAACAACAACCATAGTCCCGACA 2103

QY      148  TGCCTTCATCAGTTTTCCTCGCTTCCTCTTCCTTCCTGTCAGTTTTCGCTTTCACAGAAAGCA 207
Db      2102  ATCTTCCGGTGTTTGCTTTCCTTTTCATCTTCTGGAAGAAGTTTAGCTATACCAATCT 2043

QY      208  GA 209
Db      2042  GA 2041

RESULT 3
US-09-949-016-13499
; Sequence 13499, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13499
; LENGTH: 54245
; TYPE: DNA
; ORGANISM: Human
;
US-09-949-016-13499

Query Match          12.6%;   Score 31.6;   DB 4;   Length 54245;
Best Local Similarity 50.7%;   Pred No. 8.6;
Matches 76; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY      38  CAAATGGATTAAATGCATTGTGCAGGGAGCCGGCCAAAGCTTGAGAGCTCTTCTCGCTG 97
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      52069  CAAATAGGTTATTTGTACAGTCAGAGGGCAACAGGAGATGCAGGCTTCAAGGGCAAG 52128

QY      98  GGAGGCCCTTGGATGTGGCCCTGAAGTACCTGGCAGCGAGCTGCATGCTTTTCATC 157
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      52129  AGAGGCCACAAAGGAATATGGTGGGAGTAAAGCAACATCGTCTGCTTCATCTTTTCC 52188

QY      158  TAGTTTCCTCGCTTCCCTTTCTGCGAG 187
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      52189  TAGGCTTGGCACTGCGCTTTTCTCTTCACG 52218
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

52189 TAGGCTTGGCACTGCCTTTTCCCTTCTCAG 52218

Matches	93;	Conservative	0;	Mismatches	103;	Indels	0;	Gaps	0;
Qy	46	TTTAAATGCA	TGTGAGGAGCCG	CCAAAAGCTTTGAGAGCTCCTT	CGCTGGGAGGCC	105			
Db	93014	TTTCAGGGTGTG	GGGGGACGACGAC	CCCTAA	CAGGGGGCTCCTGCTCTATCT	CAGTCA	93073		
Qy	106	CTTGGAA	TGGCCTGAAAGTAAGCTTGGCAGGAGCCTGACATGCTTTT	CACTAGTTTCC	165				
Db	93074	CTACCTAAGCAGGAGGAAAGATGGGTGCCAGCTGGGCTTGGCGCCATGCTCTCTGTTTCA	93133						
Qy	166	TCGCTTCTCTCTCTTTCTG	CAGTCTTTTCGCTTTCACAGAAAGCAGAATCCTTTAAAAATAACC	225					
Db	93134	CCAGGGCCCTTCACTGGAGCTGTATCCCTTAGAGATATGGTAGCTCTTTTGGAGGCAGA	93193						
Qy	226	CTCTTAGTTACATCT	241						
Db	93194	ATCAGATTTTCA	CAATTT	93209					

RESULT 6  
US-09-949-016-11945/c  
; Sequence 11945, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,769  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11945  
; LENGTH: 86439  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11945

Query Match	12.4%	Score 31;	DB 4;	Length 86439;
Best Local Similarity	57.9%;	Pred. No. 19;		
Matches	55;	Conservative 0;	Mismatches 40;	Indels 0;
Gaps	0;			

  

Qy	141	CTTGACATGCTTTCATCTAGTTTCTCGCTTCTCTTCTCTTCTCGAGTTTTCGCTTCACA	200
Db	66633	CTCCCAAGTCTTGTGCTCTTCTTCTTGTGTACCATCTCTTATGAATTTTTTTTTTAAACA	66574

  

Qy	201	GAAGCAGAAATCCTTAAAAATACCCCTCTTAGTTC	235
Db	66573	GCATTGCGAGTCTCTGAAATGTCTTGAAGTTC	66539

RESULT 7  
US-09-949-016-16990/c  
; Sequence 16990, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16990
; LENGTH: 86440
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16990

Query Match          12.4%; Score 31; DB 4; Length 86440;
Best Local Similarity 57.9%; Pred No. 19;
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy      141  CCTGACATGCTTTCATCTAGTTTCCTCGCTTCCTTTCTCGAGTTTTCGCTTCACA 200
Db      66633  CCTCCCAAGTCTCTTTGTCTCTCTTTCTTTGTTACCCATCCTTTATGAATTTTTTTTTTAACA 66574

Qy      201  GAAAGCAGAATCCTTTAAAAATAAACCTCTTAGTTC 235
Db      66573  GCATTGCGAGTCTTTGAAATGTGCTTGAAGTTC 66539

```

RESULT 8  
US-09-949-016-177614/c  
; Sequence 177614, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 177614  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-177614

	Query Match	12.3%	Score 30.8;	DB 4;	Length 601;
	Best Local Similarity	56.0%;	Pred. No. 1.1;		
	Matches 56;	Conservative 1;	Mismatches 43;	Indels 0;	Gaps 0
Qy	143	TGACATGCTTTTCATCTCTAGTTTCCTCGCTCTCTCTCTTTTCGACGATTTTCGCTTCACAGA	202		
Db	334	TCAAAATGCCTGATTACACTTTCCCACTTCGCTAYCCTTTCTCTTTATTGCTCTCTCACCCAC	275		
Qy	203	AAGCAGAAATCCTTTAAAAATAACCCCTCTTAGTTTCACATCTG	242		
Db	274	AGCAAAACTCCTCAAAAATATCTATCTTTTGCAATCTTTCAG	235		

RESULT 9  
US-09-949-016-13877  
; Sequence 13877, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

us-09-899-276c-4.rni

Tue Aug 9 17:01:26 2005

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13877
; LENGTH: 202111
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(202111)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13877

Query Match          12.3%; Score 30.8; DB 4; Length 202111;
Best Local Similarity 57.1%; Pred. No. 37;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 119 CTGAAGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCGCTTCTTCT 178
Db 77594 CTGAGTGACCTCTTAAGCAAGCGACGCTGTTTGTACATTTTCTTCTCTGCT 77653

QY 179 TTTCTGCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTA 216
Db 77654 TTTGTTCAATGTTGTTTCTCAGTGAATGTCCTCTGA 77691

RESULT 10
US-09-949-016-14449
; Sequence 14449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14449
; LENGTH: 34088
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14449

Query Match          12.2%; Score 30.6; DB 4; Length 34088;
Best Local Similarity 52.8%; Pred. No. 15;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 106 CTGGGAATGCGCTGAAGTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCC 165
Db 21671 CTGGAGTGCCGCCCGAGGAGTAGTCTCAGGTATGCGTATATTTGAGAAGCTTTTGC 21730

QY 166 TCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 225
Db 21731 CCACACACAGCTTTTTCGCTTTCAGTTTTCGCTTTCACAGAAAGCAGAAATCCTTAAATAACC 225

QY 226 CTCTT 230
Db 21791 ATCTT 21795

RESULT 11
US-09-949-016-58723

```

```

; Sequence 58723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58723
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-58723

Query Match          12.1%; Score 30.2; DB 4; Length 601;
Best Local Similarity 52.8%; Pred. No. 1.8;
Matches 65; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 44 GATTTAATGCAATTCAGGGAGCGCGCAAGCTTGAGAGCTCTCTCTGCTGGGAGGC 103
Db 137 GAAACACTTCATCTGCTGTGTCAGCCAGCTCACCTTGACAGGTATCTCTGCTGGCGTAA 196

QY 104 CCCTTGAATGCGCTGAGCTGAAGCTAAGCTGGCAGCGAGCTGACATGCTTTCATCTAGTTT 163
Db 197 CCCAAGGGCCATGTGAAAGGACATTTCACTGAGGACATGATGGGTGCTTGTGTTT 256

QY 164 CCT 166
Db 257 CTT 259

RESULT 12
US-09-949-016-17188
; Sequence 17188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17188
; LENGTH: 31797
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-17188

Query Match          12.1%; Score 30.2; DB 4; Length 31797;
Best Local Similarity 53.9%; Pred. No. 20;
Matches 62; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 129 GCTGGCAGCGAGCTGACATGCTTTCATCTAGTTTCTCTGCTTCTCTCTCTCTCTCT 188
Db 19995 GCCTGTACCGGGTGTCCCCACCATAGATCTGGGATGCTCTCTCTCTCTCTCTCTG 20054

```



***This Page Blank (uspto)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 445.951 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-4  
Perfect score: 250  
Sequence: 1 aggtctctatgatgctacta.....aggtcacatctgtggtcagt 250

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues  
Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*

18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*

19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*

20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*

21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*

22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*

24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*

25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	100.0	250	9	US-09-899-276-4
2	250	100.0	681	19	US-10-283-975A-95
3	250	100.0	3221	19	US-10-685-705-3
4	250	100.0	11793	19	US-10-685-705-4
5	237.4	95.0	2776	22	US-10-833-656-2
6	162.2	64.9	5926	15	US-10-311-455-1626
7	160.4	64.2	5926	15	US-10-311-455-1625

RESULT 1

US-09-899-276-4

; Sequence 4, Application US/09899276

; Patent No. US20020106355A1

; GENERAL INFORMATION:

; APPLICANT: Rosi, Frank

; APPLICANT: Soto, Ubaldo

; APPLICANT: Coy, Johannes

; APPLICANT: Finzer, Patrick

; APPLICANT: Delius, Hajo

; APPLICANT: Poustka, Annemarie

; APPLICANT: zur Hausen, Harald

; APPLICANT: Patzelt, Andrea

; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene

; FILE REFERENCE: 012627-023

; CURRENT APPLICATION NUMBER: US/09/899,276

; PRIOR FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: EP 00 114 560.6

; PRIOR FILING DATE: 2000-07-06

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 250

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-899-276-4

Query Match 100.0%; Score 250; DB 9; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1e-72;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

Sequence 121412,  
Sequence 121412,  
Sequence 137, App  
Sequence 9824, Ap  
Sequence 572, App  
Sequence 1357, Ap  
Sequence 45154, A  
Sequence 45154, A  
Sequence 892, App  
Sequence 2837, Ap  
Sequence 19, Appl  
Sequence 19, Appl  
Sequence 15197, A  
Sequence 15197, A  
Sequence 24652, A  
Sequence 24652, A  
Sequence 1, Appli  
Sequence 695, App  
Sequence 695, App  
Sequence 743, App  
Sequence 327, App  
Sequence 325, App  
Sequence 6919, Ap  
Sequence 136073,  
Sequence 136073,  
Sequence 1372, Ap  
Sequence 7118, Ap  
Sequence 17586, A  
Sequence 221715,  
Sequence 221715,  
Sequence 203608,  
Sequence 203608,  
Sequence 18021, A  
Sequence 194, App  
Sequence 107659,  
Sequence 4338, Ap  
Sequence 6752, Ap  
Sequence 96, Appl  
Sequence 109, App

US-10-027-632-121412 13  
US-10-027-632-121412 17  
US-10-322-281-137 19  
US-09-867-701-9824 9  
US-09-822-849A-572 9  
US-10-723-860-1357 20  
US-10-027-632-45154 13  
US-10-027-632-45154 17  
US-10-087-192-892 13  
US-10-723-860-2837 20  
US-09-940-227-19 20  
US-10-933-058-19 21  
US-10-425-115-15197 20  
US-10-021-323-1950 19  
US-10-357-930-24652 20  
US-10-109-529A-1 14  
US-10-172-118-695 17  
US-10-342-887-695 18  
US-10-956-157-743 21  
US-10-295-027-327 17  
US-10-295-027-325 17  
US-10-719-993-6919 20  
US-10-027-632-136073 13  
US-10-027-632-136073 17  
US-10-087-192-1972 13  
US-10-719-993-7118 20  
US-10-741-600-17586 21  
US-10-027-632-221715 13  
US-10-027-632-221715 17  
US-10-027-632-203608 13  
US-10-027-632-203608 17  
US-10-282-122A-18021 17  
US-10-243-552-194 17  
US-10-424-599-107659 18  
US-10-956-157-4338 21  
US-10-719-993-6752 20  
US-10-175-523-96 14  
US-10-767-701-109 19

49 19.6 934 13  
49 19.6 934 17  
36.4 14.6 109559 19  
35.8 14.3 165 9  
33.6 13.4 673 9  
33.2 13.3 227968 20  
32.8 13.1 504 13  
32.8 13.1 504 17  
32.8 13.1 23406 13  
32.8 13.1 260160 20  
32.6 13.0 2092 20  
32.6 13.0 2092 21  
32.2 12.9 3772 20  
32 12.8 469 19  
31.6 12.6 1118 20  
31.6 12.6 3417 14  
31.6 12.6 5879 17  
31.6 12.6 5879 18  
31.6 12.6 5879 21  
31.6 12.6 6921 17  
31.6 12.6 7043 17  
31.4 12.6 13487 20  
31.2 12.5 819 13  
31.2 12.5 819 17  
31.2 12.5 51615 13  
31 12.4 201 20  
31 12.4 24475 21  
30.8 12.3 632 13  
30.8 12.3 632 17  
30.8 12.3 656 13  
30.8 12.3 656 17  
30.8 12.3 4491 17  
30.6 12.2 1074 17  
30.6 12.2 1086 18  
30.6 12.2 4305 21  
30.6 12.2 68744 20  
30.6 12.2 110079 14  
30.4 12.2 433 19

Tue Aug 9 17:01:26 2005

RESULT 3  
US-10-685-705-3  
; Sequence 3, Application US/10685705  
; Publication No. US20040177387A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: JAYAKRISHNA, Ambati  
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
; TITLE OF INVENTION: Degeneration  
; FILE REFERENCE: 050229-0415  
; CURRENT APPLICATION NUMBER: US/10/685,705  
; PRIOR FILING DATE: 2003-10-16  
; PRIOR FILING DATE: 2002-10-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 3221  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-685-705-3

Query Match 100.0%; Score 250; DB 19; Length 3221;  
Best Local Similarity 100.0%; Pred. No. 2.9e-72; Indels 0; Gaps 0;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 60  
Db 2247 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 2306  
QY 61 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 120  
Db 2307 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 2366  
QY 121 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 180  
Db 2367 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 2426  
QY 181 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 240  
Db 2427 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 2486  
QY 241 TGTGGTCACT 250  
Db 2487 TGTGGTCACT 2496

RESULT 4  
US-10-685-705-4  
; Sequence 4, Application US/10685705  
; Publication No. US20040177387A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: JAYAKRISHNA, Ambati  
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
; TITLE OF INVENTION: Degeneration  
; FILE REFERENCE: 050229-0415  
; CURRENT APPLICATION NUMBER: US/10/685,705  
; PRIOR FILING DATE: 2003-10-16  
; PRIOR FILING DATE: 2002-10-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 11793  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-685-705-4

Query Match 100.0%; Score 250; DB 19; Length 11793;  
Best Local Similarity 100.0%; Pred. No. 5.1e-72;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 60  
Db 1 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 60  
QY 61 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 120  
Db 61 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 120  
QY 121 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 180  
Db 121 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 180  
QY 181 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 240  
Db 181 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 240  
QY 241 TGTGGTCACT 250  
Db 241 TGTGGTCACT 250

RESULT 2  
US-10-283-975A-95  
; Sequence 95, Application US/10283975A  
; Publication No. US20040110792A1  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT APPLICATION NUMBER: US/10/283,975A  
; CURRENT FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 95  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-283-975A-95

Query Match 100.0%; Score 250; DB 19; Length 681;  
Best Local Similarity 100.0%; Pred. No. 1.5e-72;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 60  
Db 40 AGGCTTCTATGATGCTACTTATCTGATTTGATGAGCAAAATGGATTTAATGCAATGTCA 99  
QY 61 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 120  
Db 100 GGGAGCGCGCCAAAGCTTGAAGCTCTTCTGGCTGGAGGCCCTTGGAAATGGCCCT 159  
QY 121 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 180  
Db 160 GAAGTAAAGCTGGCAGCGCCCTGACATGCTTTCACTAGTTTCTCGCTTCCTTCTCTT 219  
QY 181 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 240  
Db 220 TCTGAGTTTTCGCTTCCACAGAAAGCAAAATCTTTAAATAACCCCTCTTAGTTACATC 279  
QY 241 TGTGGTCACT 250  
Db 280 TGTGGTCACT 289

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 60  
DB 7050 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 7109  
QY 61 GGGAGCGGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTCGCCT 120  
DB 7110 GGGAGCGGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTCGCCT 7169  
QY 121 GAAGGTAAGCTGGCAGCGCCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCTCTTT 180  
DB 7170 GAAGGTAAGCTGGCAGCGCCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCTCTTT 7229  
QY 181 TCTGCAGTTTTCCTTTCACAGAAAGCAGAAATCTTAAATAAACCTCTTCTAGTTTCATC 240  
DB 7230 TCTGCAGTTTTCCTTTCACAGAAAGCAGAAATCTTAAATAAACCTCTTCTAGTTTCATC 7289  
QY 241 TGTGGTCACT 250  
DB 7290 TGTGGTCACT 7299

## RESULT 5

US-10-833-656-2  
; Sequence 2, Application US/10833656  
; Publication No. US20050148507A1  
; GENERAL INFORMATION:  
; APPLICANT: Wandl, Robert  
; APPLICANT: Necina, Roman  
; APPLICANT: Doods, Henri  
; APPLICANT: Lenter, Martin  
; APPLICANT: Seidler, Randolph  
; TITLE OF INVENTION: Method for the production of an N-terminally modified  
; TITLE OF INVENTION: chemotactic factor  
; FILE REFERENCE: Case 1/1492  
; CURRENT APPLICATION NUMBER: US/10/833,656  
; CURRENT FILING DATE: 2004-04-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2776  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2582)..(2582)  
; OTHER INFORMATION: n = c, a, t or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2583)..(2583)  
; OTHER INFORMATION: n = c, a, t or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2585)..(2585)  
; OTHER INFORMATION: n = c, a, t or g  
US-10-833-656-2

Query Match 95.0%; Score 237.4; DB 22; Length 2776;  
Best Local Similarity 99.2%; Pred. No. 4.4e-68;  
Matches 249; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 60  
DB 35 AGGCTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 94  
QY 61 GGGAGCGGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTCGCCT 120  
DB 95 GGGAGCGGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTCGCCT 154  
QY 121 GAAGGTAAGCTGGCAGCGCCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCTCTTT 180  
DB 155 GAAGGTAAGCTGGCAGCGCCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCTCTTT 214  
QY 181 TCTGCAGTTTTCCTTTCACAGAAAGCAGAAATCTTAAATAAACCTCTTCTAGTTTCACAT 239

DB 215 TCTGCAGTTTTCCTTTCAGAGAAAGCAGAAATCTTAAATAAACCTCTTCTAGTTTCACAT 274  
QY 240 CTGTGGTCACT 250  
DB 275 CTGTGGTCACT 285

## RESULT 6

US-10-311-455-1626/c  
; Sequence 1626, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1626  
; LENGTH: 5926  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1626

Query Match 64.9%; Score 162.2; DB 15; Length 5926;  
Best Local Similarity 78.5%; Pred. No. 6.8e-43;  
Matches 194; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 4 CTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 63  
DB 1430 CTTCTATGATGCTACTATTCTGCAATTTGAATGAGCAAAATGGAATTAATGATTTGCA 1371  
QY 64 AGCGGCGCAAGCTTGAGAGCTCTTCTGGCTGGAGGCCCTTGGAAATGTCGCCTGAA 123  
DB 1370 AACCGACCAAACTTAAATAAACCTCTTCTAACTTAAATAAACCTTAAATAAACCTTAA 1311  
QY 124 GGTAAAGCTGGCAGCGCCTGACATGCTTTTCATCTAGTTTCTCGCTTCCTTCTCTTTCT 183  
DB 1310 AATAAACTAAACGAACTTAAATAAACCTTAAATAAACCTTAAATAAACCTTAAATAAAC 1251  
QY 184 GCAGTTTTCCTTTCACAGAAAGCAGAAATCTTAAATAAACCTTCTTCTAGTTTCACATCTGT 243  
DB 1250 ACAATTTTCCTTTCACAGAAAGCAGAAATCTTAAATAAACCTTCTTCTAGTTTCACATCTGT 1191  
QY 244 GGTCACT 250  
DB 1190 AATCAAT 1184

## RESULT 7

US-10-311-455-1625  
; Sequence 1625, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter  
; TITLE OF INVENTION: cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455

```
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1625
; LENGTH: 5926
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1625

Query Match      64.2%; Score 160.4; DB 15; Length 5926;
Best Local Similarity 77.6%; Pred. No. 2.7e-42;
Matches 194; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTAATGCAATGTCA 60
    |||||
Db 4494 AGGTTTATGATGTTATTTTGTATTTGAATGAGTAAATGGAATTAATGTTATTTA 4553
    |||||

QY 61 GGGAGCGGCAAGCTTGAGAGCTCTTCCTGGCTGGGAGGCCCTTGGAAATGGGCT 120
    |||||
Db 4554 GGGAGTCGGTTAAAGTTTGAGAGTTTTTTTGGTTGGGAGGTTTTTGGAAATGGGTTT 4613
    |||||

QY 121 GAAGTAAGTCGACGCGCCGACATGCTTTCATCTAGTTTCCTCGCTTCCTCTCTTT 180
    |||||
Db 4614 GAAGGTAAGTTGGTAGCGATTTTGATATGTTTTTATTTAGTTTTCGTTTTTTTTTT 4673
    |||||

QY 181 TCTGCAAGTTTCGCTTCCACAGAAAGCAGATCCTTAAATAACCCCTCTTAGTTCACATC 240
    |||||
Db 4674 TTGTAGTTTCGTTTAGAGAAAGTAGAATTTTAAATAATATTTTATTTAGTTTATTT 4733
    |||||

QY 241 TGTGGTCAGT 250
    |||||
Db 4734 TGTGGTTAGT 4743
    |||||
```

```
RESULT 8
US-10-027-632-121412
; Sequence 121412, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121412
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-121412
```

```
Query Match      19.6%; Score 49; DB 13; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTA 49
    |||||
Db 886 AGGCTTCTATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTA 934
    |||||

RESULT 9
US-10-027-632-121412
; Sequence 121412, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121412
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-121412

Query Match      19.6%; Score 49; DB 17; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCTTCTATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTA 49
    |||||
Db 886 AGGCTTCTATGCTACTATTCTGCATTTGAATGAGCAAAATGGAATTA 934
    |||||

RESULT 10
US-10-322-281-137/c
; Sequence 137, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 137
; LENGTH: 109559
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(109559)
; OTHER INFORMATION: n = A,T,C or G
```

US-10-322-281-137

Query Match 14.6%; Score 36.4; DB 19; Length 109559;  
Best Local Similarity 55.6%; Pred. No. 1.9; Indels 1; Gaps 1;  
Matches 90; Conservative 0; Mismatches 71; Indels 1; Gaps 1;  
QY 36 AGCAATGGATTAAATGATGTCAGGAGCGCCGCAAGCTTTGAGAGCTTCCTTCCTGGC 95  
Db 106445 AGGAGCTTGCTCGCATGTTGTGAAGGAGCCAGGTTGTCCAGGCTGCTTCGAGGC 106386  
QY 96 TGGGAGGCCCTTGGATGTGGCTGAAGTAAGCTGGGAGCGAGCTGACATGCTTTCA 155  
Db 106385 TGGGAGGAGCTAGGCAT-TGGGATGGAGCGAGGTTGTAGGCTTACCTAGCTCCATTTCT 106327  
QY 156 TCTAGTTTCCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 197  
Db 106326 TCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 106285

RESULT 11

US-09-867-701-9824/c  
; Sequence 9824, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9824  
; LENGTH: 165  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(165)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-9824

Query Match 14.3%; Score 35.8; DB 9; Length 165;  
Best Local Similarity 53.7%; Pred. No. 0.2;  
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 87 CTTCTCGCTGGGAGGCCCTTGGAAATGTGGCTGAAGCTGGCAGCGAGCTGAC 146  
Db 145 CTGCTGCGAGGTGTGGCCCTTGGCTGGACCTGGGCTGAATTTGGGAAGGTGGT 86  
QY 147 ATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 206  
Db 85 TTCTTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 86  
QY 207 AGAATCCTTAAATA 222  
Db 25 AGAATTAATAACNCTA 10

RESULT 12

US-09-822-849A-572  
; Sequence 572, Application US/09822849A  
; Patent No. US20020045170A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakur

; APPLICANT: Graham, James R.  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6403  
; CURRENT APPLICATION NUMBER: US/09/822,849A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/195,582  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 598  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 572  
; LENGTH: 673  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-849A-572

Query Match 13.4%; Score 33.6; DB 9; Length 673;  
Best Local Similarity 52.9%; Pred. No. 1.9;  
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 87 CTTCTCGCTGGGAGGCCCTTGGAAATGTGGCTGAAGCTGGCAGCGAGCTGAC 146  
Db 532 CTGCTGCGAGGTGTGGCCCTTGGCTGGACCTGGGCTGAATTTGGGAAGGTGGT 591  
QY 147 ATGCTTTCATCTAGTTTCCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 206  
Db 592 TTCTTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 651  
QY 207 AGAATCCTTAAATA 222  
Db 652 AGAATTAATAACACTA 667

RESULT 13

US-10-723-860-1357  
; Sequence 1357, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Nataasha  
; APPLICANT: Ginsburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723,860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1357  
; LENGTH: 227968  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-1357

Query Match 13.3%; Score 33.2; DB 20; Length 227968;  
Best Local Similarity 57.8%; Pred. No. 30;  
Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
QY 30 TGAATGAGCAATGGATTAAATGATGTCAGGAGCGCGCCCAAGCTTGAGAGCTTCCTT 89  
Db 218073 TGACTAGGAATTCCTTAACCTGCTTATCTGTCAATGGCCCAATTTAAGGTTCTT 218132  
QY 90 CTTGCTGGGAGGCCCTTGGAAATGTGGCTGAGCTGAAGCTAAGCT 131  
Db 218133 CCAGGCCCTCAAAACTCTATGATGATGATGATGATGATGATGATGATGATGATG 218174  
RESULT 14  
US-10-027-632-45154  
; Sequence 45154, Application US/10027632  
; Publication No. US20020198371A1

GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45154  
; LENGTH: 504  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-45154

Query Match 13.1%; Score 32.8; DB 13; Length 504;  
Best Local Similarity 48.4%; Pred. No. 3.2;  
Matches 88; Conservative 1; Mismatches 93; Indels 0; Gaps 0;  
QY 18 CTATTCGCAATTTGAATGAGCAAAATGATTAATGCAATGTCAGGGAGCGGCCAAAGCT 77  
DB 61 CTAGCTAAATAATTATAGTCTCATTAATTTTAAAGGATTTCTCTGTATTTCTAGCAT 120  
QY 78 TGAGAGCTCCTTCCTCGCTGGGAGGCCCTTGGAATGTGGCCCTGAAGCTGAAGCTGGCAGC 137  
DB 121 TTAATATTGCTGATAATGTCGATGCCAGTCTGATTTTGTCTTTTGTAAAGATAAAGAA 180  
QY 138 GAGCCTGACATGCTTTCATCTAGTTTCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 197  
DB 181 ATATTTTATCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 240  
QY 198 AC 199  
DB 241 TC 242

RESULT 15  
US-10-027-632-45154  
; Sequence 45154, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45154  
; LENGTH: 504  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-45154  
Query Match 13.1%; Score 32.8; DB 17; Length 504;  
Best Local Similarity 48.4%; Pred. No. 3.2;  
Matches 88; Conservative 1; Mismatches 93; Indels 0; Gaps 0;  
QY 18 CTATTCGCAATTTGAATGAGCAAAATGATTAATGCAATGTCAGGGAGCGGCCAAAGCT 77  
DB 61 CTAGCTAAATAATTATAGTCTCATTAATTTTAAAGGATTTCTCTGTATTTCTAGCAT 120  
QY 78 TGAGAGCTCCTTCCTCGCTGGGAGGCCCTTGGAATGTGGCCCTGAAGCTGAAGCTGGCAGC 137  
DB 121 TTAATATTGCTGATAATGTCGATGCCAGTCTGATTTTGTCTTTTGTAAAGATAAAGAA 180  
QY 138 GAGCCTGACATGCTTTCATCTAGTTTCTCGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 197  
DB 181 ATATTTTATCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 240  
QY 198 AC 199  
DB 241 TC 242

Search completed: August 4, 2005, 15:54:16  
Job time : 448.951 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 1205.83 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-5  
Perfect score: 300  
Sequence: 1 aaggaggaggcagtggtta.....ataaccaggatgaacttct 300

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	300	100.0	300	6	AX343332	AX343332 Sequence
2	300	100.0	3227	9	HUMMCP1	D26087 Human gene
3	298.4	99.5	9174	9	AF519531	AF519531 Homo sapi
4	298.4	99.5	147416	9	AC005549	AC005549 Homo sapi
5	296.8	98.9	3221	9	AY357296	AY357296 Homo sapi
6	296.8	98.9	11793	9	HSY18933	Y18933 Homo sapien
7	284.8	94.9	73806	2	AC021520	AC021520 Homo sapi
8	204	68.0	5926	6	AX346554	AX346554 Sequence
9	194.4	64.8	5926	6	AX346555	AX346555 Sequence
10	87.6	29.2	3405	4	BOVMCP1X	L32659 Bovine mono
11	68	22.7	2788	10	MMU12470	U12470 Mus musculu
12	68	22.7	160869	2	AC022299	AC022299 Mus muscu
13	68	22.7	180944	2	AC073824	AC073824 Mus muscu
14	68	22.7	222121	10	AC012294	AC012294 Mus muscu
15	68	22.7	223726	10	AL626807	AL626807 Mouse DNA
16	68	22.7	240425	10	AL713839	AL713839 Mus muscu
17	66.4	22.1	28613	10	MM0238892	AJ238892 Mus muscu
18	54.6	18.2	3657	10	AF079313	AF079313 Rattus no
19	54.6	18.2	242335	2	AC123203	AC123203 Rattus no

20	54.6	18.2	245808	2	AC114440	AC114440 Rattus no
21	40.4	13.5	181022	9	AC011193	AC011193 Homo sapi
22	40	13.3	67962	9	AL596113	AL596113 Human DNA
23	40	13.3	178414	2	AC025341	AC025341 Homo sapi
24	40	13.3	198190	2	AC136446	AC136446 Homo sapi
25	39.2	13.1	153053	9	AC007533	AC007533 Homo sapi
26	39	13.0	129660	9	AC130814	AC130814 Homo sapi
27	39	13.0	159471	9	AC092594	AC092594 Homo sapi
28	39	13.0	234795	10	AC099934	AC099934 Mus muscu
29	38.6	12.9	147772	5	BX649242	BX649242 Zebrafish
30	38.6	12.9	229319	2	CR354585	CR354585 Danio rer
31	38	12.7	261369	2	AC106158	AC106158 Rattus no
32	37.6	12.5	140616	9	AC127389	AC127389 Homo sapi
33	37.4	12.5	170743	2	AC102789	AC102789 Mus muscu
34	37.2	12.4	183646	9	AC007339	AC007339 Homo sapi
35	37.2	12.4	185430	2	AC025706	AC025706 Homo sapi
36	37.2	12.4	187378	9	AL357150	AL357150 Human DNA
37	37.2	12.4	248750	10	AC112685	AC112685 Mus muscu
38	37	12.3	958	3	NAM556126	AJ556126 Necator a
39	37	12.3	958	3	NAM556131	AJ556131 Necator a
40	37	12.3	959	3	NAM556125	AJ556125 Necator a
41	37	12.3	959	3	NAM556128	AJ556128 Necator a
42	37	12.3	959	3	NAM556130	AJ556130 Necator a
43	37	12.3	960	3	NAM556124	AJ556124 Necator a
44	37	12.3	13605	3	NAM417719	AJ417719 Necator a
45	37	12.3	181021	10	AL663031	AL663031 Mouse DNA

ALIGNMENTS

RESULT 1  
AX343332  
LOCUS AX343332  
DEFINITION Sequence 5 from Patent EP1170372.  
ACCESSION AX343332  
VERSION AX343332.1 GI:18491682  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Roesl,F., Soto,U., Coy,J., Finzer,P., Delius,H., Poustka,A., zur Hausen,H. and Patzelt,A.  
TITLE Regulatory sequences of the human mcp-1 gene  
JOURNAL Patent: EP 1170372-A 5 09-JAN-2002;  
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS (DE)  
FEATURES  
source Location/Qualifiers  
1..300  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

/wd_xref=taxon:9606					
ORIGIN	Query Match	100.0%	Score 300;	DB 6;	Length 300;
	Best Local Similarity	100.0%;	Pred. No. 4e-83;		
	Matches 300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	AAGGAGGAGCGATGGGCTAGGAGATCGAGAGATCAGAAATTTTAACTCAGCCGAGCCA	60		
Db	1	AAGGAGGAGCGATGGGCTAGGAGATCGAGAGATCAGAAATTTTAACTCAGCCGAGCCA	60		
Qy	61	TTAACTATGCTCCTCAAGTACTCCTATCATATTTGTAGAGACAACAGTTCTCACTGAAATGAAT	120		
Db	61	TTAACTATGCTCCTCAAGTACTCCTATCATATTTGTAGAGACAACAGTTCTCACTGAAATGAAT	120		
Qy	121	TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC	180		
Db	121	TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC	180		
Qy	181	AACTGAGGAATGAAGTCAGGCTTTTCCAAATTCGCGAAATACTCTCTCCACTGCTTACTTCATG	240		

```

/product="monocyte chemoattractant protein-1"
/protein_id="BAA05080.1"
/db_xref="GI:992947"
/translation="MKVSAALLCLLLIAATFIPOGLAQP"
3148..>3227
/number=1

intron
ORIGIN
Query Match 100.0%; Score 300; DB 9; Length 3227;
Best Local Similarity 100.0%; Pred. No. 4.4e-83;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGAGGAGCGAGTGGCTAGGAGATCGAGAGATCAGAAATTTAAACTCAGCCGACCA 60
Db 1712 AAGGAGGAGCGAGTGGCTAGGAGATCGAGAGATCAGAAATTTAAACTCAGCCGACCA 1771
QY 61 TTAACATGCTCAAGTACTCTTATCATATTTGTAAGAGACAACAGTTCACTGAATGAAT 120
Db 1772 TTAACATGCTCAAGTACTCTTATCATATTTGTAAGAGACAACAGTTCACTGAATGAAT 1831
QY 121 TCTAAGGTCCTTGGGTTTTTATCAGTGCTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180
Db 1832 TCTAAGGTCCTTGGGTTTTTATCAGTGCTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 1891
QY 181 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCGAAATATCTCTCCACTGCTTACTCATG 240
Db 1892 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCGAAATATCTCTCCACTGCTTACTCATG 1951
QY 241 TCCCTTGGAAATTAAGAGAGCGGAGCATAGTGCATACACAGGAGTGAATCTTCT 300
Db 1952 TCCCTTGGAAATTAAGAGAGCGGAGCATAGTGCATACACAGGAGTGAATCTTCT 2011

RESULT 3
AF519531 9174 bp DNA linear PRI 09-JUL-2002
LOCUS Homo sapiens small inducible cytokine A2 (monocyte chemotactic
DEFINITION protein 1) (SCYA2) gene, complete cds.
ACCESSION AF519531
VERSION AF519531.1 GI:21435976
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9174)
AUTHORS Rieder,M.J., Armel,T.Z., Carrington,D.P., Ozuna,M., KuldaneK,S.A.,
Rajkumar,N., Toth,E.J., Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattleSNPs. NHLBI HL66682 Program
for Genomic Applications, UW-PHCRC, Seattle, WA (URL:
http://pga.gs.washington.edu).
FEATURES
source
1..9174
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
variation
107
/replaces="g"
/replace="g"
variation
291
/frequency="0.03"
/replace="g"
variation
305
/frequency="0.15"
/replace="t"
variation
334
/frequency="0.01"
/replace="g"
variation
604
/frequency="0.16"

181 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCGAAATACTCTCCACTGCTTACTCATG 240
241 TCCCTTGGAAATTAAGAGAGCGGAGCATAGTGCATACACAGGAGTGAATCTTCT 300
241 TCCCTTGGAAATTAAGAGAGCGGAGCATAGTGCATACACAGGAGTGAATCTTCT 300

HUMMCP1 3227 bp DNA linear PRI 14-APR-2000
LOCUS Human gene for monocyte chemoattractant protein-1, 5'-genome
DEFINITION
ACCESSION D26087
VERSION MCP-1; monocyte chemoattractant protein-1.
KEYWORDS MCP-1; monocyte chemoattractant protein-1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Rollins,B.J., Stier,P., Ernst,T. and Wong,G.G.
TITLE The human homolog of the JE gene encodes a monocyte secretory
protein
JOURNAL Mol. Cell. Biol. 9 (11), 4687-4695 (1989)
MEDLINE 90097880
PUBMED 2513477
REFERENCE 2 (sites)
AUTHORS Shyu,Y.J., Li,Y.S. and Kolattukudy,P.E.
TITLE Structure of human monocyte chemotactic protein gene and its
regulation by TPA
JOURNAL Biochem. Biophys. Res. Commun. 169 (2), 346-351 (1990)
MEDLINE 90290466
PUBMED 2357211
REFERENCE 3 (bases 1 to 3227)
AUTHORS Ueda,A., Okuda,K., Ohno,S., Shirai,A., Igarashi,T., Matsunaga,K.,
Fukushima,J., Kawamoto,S., Ishigatsubo,Y. and Okubo,T.
TITLE NF-kappa B and Sp1 regulate transcription of the human monocyte
chemoattractant protein-1 gene
JOURNAL J. Immunol. 153 (5), 2052-2063 (1994)
MEDLINE 94327939
PUBMED 8051410
REFERENCE 4 (bases 1 to 3227)
AUTHORS Ueda,A.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-1993) Atsuhisa Ueda, Yokohama City University
School of Medicine, First Department of Internal Medicine; 3-9
Fukuura, Kanazawa-ku, Yokohama 236, Japan (Tel:045-787-2630,
Fax:045-786-3444)
COMMENT Submitted (06-Dec-1993) to DDBJ by:
Atsuhisa Ueda
First Department of Internal Medicine
Yokohama City University School of Medicine
3-9 Fukuura, Kanazawa-ku
Yokohama 236
Japan
Phone: 045-787-2630
Fax: 045-786-3444.
FEATURES
source
1..3227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/tissue_type="placenta"
/clone_lib="EMBL3 SF6/T7"
400..407
/feature="kappa B"
enhancer
2948..2953
GC_signal
2975..2979
TATA_signal
3012..3147
exon
/number=1
3072..>3147
/codon_start=1

```

```

variation      /replace="t"
609
/refrequency="0.02"
/repeat_region /replace="a"
614. .720
/rpt_family="MIR"
/rpt_type=dispersed
678
/refrequency="0.14"
/repeat_region /replace="g"
1051
/refrequency="0.01"
/repeat_region /replace="a"
1054
/refrequency="0.01"
/repeat_region /replace="t"
1072
/refrequency="0.03"
/repeat_region /replace="t"
1280
/refrequency="0.23"
/repeat_region /replace="a"
1537
/refrequency="0.15"
/repeat_region /replace="g"
1548
/refrequency="0.08"
/repeat_region /replace="a"
1563
/refrequency="0.04"
/repeat_region /replace="g"
1707
/refrequency="0.02"
/repeat_region /replace="c"
1718. .1830
/rpt_family="MIR"
/rpt_type=dispersed
1729
/refrequency="0.01"
/repeat_region /replace="c"
1980
/refrequency="0.13"
/repeat_region /replace="t"
1984
/refrequency="0.02"
/repeat_region /replace="g"
2307
/refrequency="0.04"
/repeat_region /replace="g"
2391
/refrequency="0.01"
/repeat_region /replace="a"
2569
/refrequency="0.08"
/repeat_region /replace="t"
3001
/refrequency="0.03"
/repeat_region /replace="a"
3190
/refrequency="0.14"
/repeat_region /replace="c"
3228. .3445
/rpt_family="MIR"
/rpt_type=dispersed
3719
/refrequency="0.01"
/repeat_region /replace="a"
3756
/refrequency="0.25"
/repeat_region /replace="c"
4065. .5981
/gene="SCYA2"
mRNA      join(4065. .4193,4990. .5107,5490. .5981)

```

---

```

/gene="SCYA2"
/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
join(4118. .4193,4990. .5107,5490. .5595)
/gene="SCYA2"
/codon_start=1
/product="small inducible cytokine A2 (monocyte
chemotactic protein 1)"
/protein_id="AA054046.1"
/db_xref="GI:21435977"
/translation="MKVSAALLCLLLIAATFIPOGLAOPDAINAPVTCYNFTNRKIS
VORLASVRRITSSKCPKEAVIFKTIKAIKADPKQKWQDSMDHLDKQOTPKT"
4296. .4409
/rpt_family="MIR"
/rpt_type=dispersed
4359. .4462
/gene="SCYA2"
/note="Region not scanned for variation"
4748
/gene="SCYA2"
/refrequency="0.26"
/repeat_region /replace="n"
4777. .4814
/rpt_family="MIR"
/rpt_type=dispersed
4881
/gene="SCYA2"
/refrequency="0.14"
/repeat_region /replace="g"
5018
/gene="SCYA2"
/refrequency="0.47"
/repeat_region /replace="c"
5132
/gene="SCYA2"
/refrequency="0.01"
/repeat_region /replace="t"
5660
/gene="SCYA2"
/refrequency="0.18"
/repeat_region /replace="t"
6184
/refrequency="0.01"
/repeat_region /replace="c"
6518
/refrequency="0.01"
/repeat_region /replace="t"
6531
/refrequency="0.05"
/repeat_region /replace="t"
6837. .6916
/rpt_family="L2"
/rpt_type=dispersed
7436
/refrequency="0.07"
/repeat_region /replace="t"
7605
/refrequency="0.03"
/repeat_region /replace="a"
7614
/refrequency="0.01"
/repeat_region /replace="a"
7659
/refrequency="0.03"
/repeat_region /replace="a"
7710
/refrequency="0.05"
/repeat_region /replace="a"
7843
/refrequency="0.31"
/repeat_region /replace="c"
8072
/refrequency="0.01"

```



repeat\_region /rpt\_family="MLT11"  
complement(5430..5549)  
repeat\_region /rpt\_family="FLAM C"  
complement(5941..6051)  
repeat\_region /rpt\_family="MER81"  
6131..6250  
repeat\_region /rpt\_family="(TA)n"  
6406..6455  
repeat\_region /rpt\_family="AT\_rich"  
6574..6600  
repeat\_region /rpt\_family="MLT1D"  
6601..6880  
repeat\_region /rpt\_family="AluJb"  
6881..6935  
repeat\_region /rpt\_family="MLT1D"  
7119..7360  
repeat\_region /rpt\_family="MLT1D"  
7337..7636  
repeat\_region /rpt\_family="MLT1D"  
7689..7888  
repeat\_region /rpt\_family="L1ME3"  
8023..8048  
repeat\_region /rpt\_family="AT\_rich"  
8871..8986  
repeat\_region /rpt\_family="MIR"  
10109..10258  
repeat\_region /rpt\_family="MER5A"  
complement(11031..11072)  
repeat\_region /rpt\_family="MIR"  
11567..11836  
repeat\_region /rpt\_family="L2"  
12501..12981  
repeat\_region /rpt\_family="MER9"  
complement(12983..13257)  
repeat\_region /rpt\_family="AluJb"  
15248..15364  
repeat\_region /rpt\_family="MIR"  
complement(16483..16653)  
repeat\_region /rpt\_family="MIR"  
16830..16949  
repeat\_region /rpt\_family="MER5A"  
17203..17380  
repeat\_region /rpt\_family="LTR16A"  
17393..18306  
repeat\_region /rpt\_family="HERVL"  
18315..24307  
repeat\_region /rpt\_family="LIP2A"  
24309..24335  
repeat\_region /rpt\_family="AT\_rich"  
24513..25245  
repeat\_region /rpt\_family="HERVL"  
25613..25758  
repeat\_region /rpt\_family="HERVL"  
26135..26228  
repeat\_region /rpt\_family="LTR16A"  
complement(27814..27950)  
repeat\_region /rpt\_family="MER94"  
28930..28968  
repeat\_region /rpt\_family="AT\_rich"  
29960..30341  
repeat\_region /rpt\_family="MLT11"  
32364..32581  
repeat\_region /rpt\_family="MIR"  
32601..32902  
repeat\_region /rpt\_family="AluSx"  
complement(33593..39359)  
repeat\_region /rpt\_family="LIP2A"  
39362..39414  
repeat\_region /rpt\_family="A-rich"  
complement(39920..40047)  
repeat\_region /rpt\_family="MIR"  
40112..40151  
repeat\_region /rpt\_family="AT\_rich"

repeat\_region complement(40152..40442)  
/rpt\_family="AluJo"  
repeat\_region complement(40810..40898)  
/rpt\_family="MLT1J"  
41293..41469  
repeat\_region /rpt\_family="LIP2A"  
complement(41846..41956)  
/rpt\_family="purine-rich"  
repeat\_region complement(42486..42569)  
/rpt\_family="MLT1I"  
42975..43098  
repeat\_region /rpt\_family="MER5A"  
43554..43983  
repeat\_region /rpt\_family="LTR24"  
44781..44811  
repeat\_region /rpt\_family="AT\_rich"  
46184..46296  
repeat\_region /rpt\_family="MIR"  
complement(47075..47140)  
repeat\_region /rpt\_family="MLT1G"  
complement(47266..47307)  
/rpt\_family="(CA)n"  
repeat\_region complement(47363..47442)  
/rpt\_family="MLT1G"  
complement(48165..48209)  
Query Match 99.5%; Score 298.4; DB 9; Length 147416;  
Best Local Similarity 99.7%; Pred. No. 1.6e-82;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AAGGAGGAGCGAGTGGCTAGGAGATCGAGAGATCAGAAATTTAAACTCAGCCAGCCA 60  
DB 62778 AAGGAGGAGCGAGTGGCTAGGAGATCGAGAGATCAGAAATTTAAACTCAGCCAGCCA 62719  
QY 61 TTAACATGCTCAAGTACTCTCATATTTTGAAGAGACACAGTTTCACTGAATGAAT 120  
DB 62718 TTAACATGCTCAAGTACTCTCATATTTTGAAGAGACACAGTTTCACTGAATGAAT 62659  
QY 121 TCTAAGGCTTTGGGTTTTATCAGTGCTCTTCTGAGTTTCTGAGGAATCTAAGGCAC 180  
DB 62658 TCTAAGGCTTTGGGTTTTATCAGTGCTCTTCTGAGTTTCTGAGGAATCTAAGGCAC 62599  
QY 181 AACTGAGGAATCAAGTCAGGCTTTCCAATTCGGAATACTCCTCCACTGCTTACTCATG 240  
DB 62598 AACTGAGGAATCAAGTCAGGCTTTCCAATTCGGAATACTCCTCCACTGCTTACTCATG 62539  
QY 241 TCCCTTGGAAATTAAGAGGAGCCAGGAGTAGCTGCCATAACACAGGATGAATCTCT 300  
DB 62538 TCCCTTGGAAATTAAGAGGAGCCAGGAGTAGCTGCCATAACACAGGATGAATCTCT 62479  
RESULT 5  
LOCUS AY357296 3221 bp DNA linear PRI 15-SBP-2003  
DEFINITION Homo sapiens MCP1 (MCP1) gene, promoter region and partial cds.  
ACCESSION AY357296  
VERSION AY357296.1 GI:34559719  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3221)  
AUTHORS Nyquist, P.A. and Degraaba, T.J.  
TITLE Two New Single Nucleotide Polymorphisms (SNPs) in the MCP-1  
Promoter in Patients with Carotid Atherosclerosis: Transcriptional  
Induction and New Protein Binding Sites  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3221)  
AUTHORS Nyquist, P.A. and Degraaba, T.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-2003) Medicine, Inova Fairfax, 3300 Gallows Road,  
Falls Church, VA 22402-3300, USA

FEATURES	Location/Qualifiers	
	1..3221	/organism="Homo sapiens"
Source	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
gene	225..>3221	
	/gene="MCP1"	
promoter	/note="HSMCR30; MCP-1"	
	225..2746	
variation	/gene="MCP1"	
	225	
variation	/gene="MCP1"	
	/note="SNP"	
variation	/replace="c"	
	668	
variation	/gene="MCP1"	
	/note="SNP"	
variation	/replace="t"	
	1878	
variation	/gene="MCP1"	
	/note="SNP"	
variation	/replace="c"	
	2445	
TATA_signal	/gene="MCP1"	
	/note="SNP"	
TATA_signal	/replace="g"	
	2710..2714	
mRNA	/gene="MCP1"	
	/evidence=not_experimental	
CDS	2747..>2882	
	/gene="MCP1"	
CDS	/product="MCP1"	
	/evidence=not_experimental	
CDS	2807..>2882	
	/gene="MCP1"	
CDS	/codon_start=1	
	/evidence=not_experimental	
CDS	/product="MCP1"	
	/protein_id="AAQ7526.1"	
CDS	/db_xref="GI:34559720"	
	/translation="MKVSAALLCLLLIAATFIQGLAQP"	
ORIGIN	Query Match 98.9%; Score 296.8; DB 9; Length 3221;	
	Best Local Similarity 99.3%; Pred. No. 4.4e-82;	
ORIGIN	Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
	Qy 1 AAGGAGGAGGCGTGGGCTAGGAGAAATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 60	
ORIGIN	Db 1445 AAGGAGGAGGCGTGGGCTAGGAGAAATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 1504	
	Qy 61 TTAACATGCCTCAAGTACTCCTATCATATTTGTGAAGACAAACAGTTCACGAAATGAAT 120	
ORIGIN	Db 1505 TTAACATGCCTCAAGTACTCCTATCATATTTGTGAAGACAAACAGTTCACGAAATGAAT 1564	
	Qy 121 TCTAAGGCTCTTGGGTTTATCATGTGCTTCTGTAGTTTCTGAGAAATCTAAGGCAC 180	
ORIGIN	Db 1565 TCTAAGGCTCTTGGGTTTATCATGTGCTTCTGTAGTTTCTGAGAAATCTAAGGCAC 1624	
	Qy 181 AACTGAGGAATGAAGTCAAGGCTTTCCAAATCCCGAAATACCTCCCTCACTGCTTACTCATG 240	
ORIGIN	Db 1625 AACTGAGGAATGAAGTCAAGGCTTTCCAAATCCCGAAATACCTCCCTCACTGCTTACTCATG 1684	
	Qy 241 TCCCTTGGAAATTAAGAGGAAGCCAGGAGATAGCTGCCATACCCAGGAGTGAACCTCT 300	
ORIGIN	Db 1685 TCCCTTGGAAATTAAGAGGAAGCCAGGAGATAGCTGCCATACCCAGGAGTGAACCTCT 1744	
	RESULT 6	
ORIGIN	HSY18933	
	LOCUS HSY18933 11793 bp DNA linear PRI 18-OCT-2000	
ORIGIN	DEFINITION Homo sapiens MCP-1 gene and enhancer region.	
	ACCESSION Y18933	
ORIGIN	Query Match 98.9%; Score 296.8; DB 9; Length 11793;	
	Best Local Similarity 99.3%; Pred. No. 4.6e-82;	
ORIGIN	Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
	Qy 1 AAGGAGGAGGCGTGGGCTAGGAGAAATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 60	
ORIGIN	Db 6249 AAGGAGGAGGCGTGGGCTAGGAGAAATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 6308	

QY 61 TTAACATGCTCAAGTACTCTATCATATTTGTAAGAGACAACTGTTCACTGAAATGAAT 120  
 Db TTAACATGCTCAAGTACTCTATCATATTTGTAAGAGACAACTGTTCACTGAAATGAAT 6368  
 QY 121 TCTAAGTCTTTGGGTTTATCAGTGCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 180  
 Db TCTAAGTCTTTGGGTTTATCAGTGCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 6428  
 QY 181 AACTGAGGAATGAAGTCAAGCTTTCCAAATTCCTCGAATACTCTCCACTGCTTACTCATG 240  
 Db AACTGAGGAATGAAGTCAAGCTTTCCAAATTCCTCGAATACTCTCCACTGCTTACTCATG 6488  
 QY 241 TCCCTTGGAAATTAAGAGGAAGCCAGGACATAGCTGCCATACCAAGGATGAATCTTCT 300  
 Db TCCCATGGAATTAAGAGGAAGCCAGGACATAGCTGCCATACCAAGGATGAATCTTCT 6548

## RESULT 7

AC021520/c

LOCUS

DEFINITION

Homo sapiens chromosome 17 clone RP11-329H16 map 17, LOW-PASS

SEQUENCE SAMPLING.

AC021520

VERSION

HTG; HTGS PHASE0.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 73806)

Birken, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE

Homo sapiens chromosome 17, clone RP11-329H16

JOURNAL

Unpublished

2 (bases 1 to 73806)

Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Becker, R., Bedalov, F.,

Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,

Chapel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,

DeArelano, K., Dewar, K., Domino, M., Doyle, M., Penestor, J.,

Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galegan, J.,

Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,

Lander, S., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,

Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,

McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,

Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,

Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,

Roy, A., Santos, R., Sever, P., Spencer, B., Stange-Thomann, N.,

Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,

Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705580.

All repeats were identified using RepeatMasker:

Smit, A.F.A. &amp; Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project information

Center project name: L5499

Center clone name: 329\_H16

-----

\* NOTE: This record contains 76 individual

\* sequencing reads that have not been assembled into

\* contigs. Runs of N are used to separate the reads

\* and the order in which they appear is completely

\* arbitrary. Low-pass sequencing is useful for

\* identifying clones that may be gene-rich and allows

\* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 862: contig of 862 bp in length  
 \* 863 962: gap of 100 bp  
 \* 963 1850: contig of 888 bp in length  
 \* 1851 1950: gap of 100 bp  
 \* 1951 2817: contig of 867 bp in length  
 \* 2818 2917: gap of 100 bp  
 \* 2918 3788: contig of 871 bp in length  
 \* 3789 3888: gap of 100 bp  
 \* 3889 4760: contig of 872 bp in length  
 \* 4761 4860: gap of 100 bp  
 \* 4861 5725: contig of 865 bp in length  
 \* 5726 5825: gap of 100 bp  
 \* 5826 6721: contig of 895 bp in length  
 \* 6721 6820: gap of 100 bp  
 \* 6821 7680: contig of 860 bp in length  
 \* 7681 7780: gap of 100 bp  
 \* 7781 8666: contig of 886 bp in length  
 \* 8667 8766: gap of 100 bp  
 \* 8767 9637: contig of 871 bp in length  
 \* 9638 9737: gap of 100 bp  
 \* 9738 10614: contig of 877 bp in length  
 \* 10615 10714: gap of 100 bp  
 \* 10715 11607: contig of 893 bp in length  
 \* 11608 11707: gap of 100 bp  
 \* 11708 12583: contig of 876 bp in length  
 \* 12584 12684: gap of 100 bp  
 \* 12685 13570: contig of 887 bp in length  
 \* 13571 13670: gap of 100 bp  
 \* 13671 14560: contig of 890 bp in length  
 \* 14561 14660: gap of 100 bp  
 \* 14661 15536: contig of 876 bp in length  
 \* 15537 15636: gap of 100 bp  
 \* 15637 16521: contig of 885 bp in length  
 \* 16522 16821: gap of 100 bp  
 \* 16822 17518: contig of 897 bp in length  
 \* 17519 17618: gap of 100 bp  
 \* 17619 18516: contig of 898 bp in length  
 \* 18517 18616: gap of 100 bp  
 \* 18617 19500: contig of 884 bp in length  
 \* 19501 19600: gap of 100 bp  
 \* 19601 20455: contig of 855 bp in length  
 \* 20456 20555: gap of 100 bp  
 \* 20556 21441: contig of 886 bp in length  
 \* 21442 21541: gap of 100 bp  
 \* 21542 22418: contig of 877 bp in length  
 \* 22419 22518: gap of 100 bp  
 \* 22519 23376: contig of 858 bp in length  
 \* 23377 23476: gap of 100 bp  
 \* 23477 24342: contig of 866 bp in length  
 \* 24343 24442: gap of 100 bp  
 \* 24443 25233: contig of 791 bp in length  
 \* 25234 25333: gap of 100 bp  
 \* 25334 26223: contig of 890 bp in length  
 \* 26224 26323: gap of 100 bp  
 \* 26324 27194: contig of 871 bp in length  
 \* 27195 27294: gap of 100 bp  
 \* 27295 28162: contig of 868 bp in length  
 \* 28163 28262: gap of 100 bp  
 \* 28263 29148: contig of 886 bp in length  
 \* 29149 29248: gap of 100 bp  
 \* 29249 30154: contig of 906 bp in length  
 \* 30155 30254: gap of 100 bp  
 \* 30255 31124: contig of 870 bp in length  
 \* 31125 31224: gap of 100 bp  
 \* 31225 32105: contig of 881 bp in length  
 \* 32106 32205: gap of 100 bp  
 \* 32206 33093: contig of 888 bp in length  
 \* 33094 33193: gap of 100 bp

* 33194	34078:	contig of 885 bp	in length
* 34079	34178:	gap of 100 bp	
* 34179	34826:	contig of 648 bp	in length
* 34827	34926:	gap of 100 bp	
* 34927	35771:	contig of 845 bp	in length
* 35772	35871:	gap of 100 bp	
* 35872	36751:	contig of 880 bp	in length
* 36752	36851:	gap of 100 bp	
* 36852	37741:	contig of 890 bp	in length
* 37742	37841:	gap of 100 bp	
* 37842	38705:	contig of 864 bp	in length
* 38706	38805:	gap of 100 bp	
* 38806	39678:	contig of 873 bp	in length
* 39679	39778:	gap of 100 bp	
* 39779	40673:	contig of 895 bp	in length
* 40674	40773:	gap of 100 bp	
* 40774	41671:	contig of 898 bp	in length
* 41672	41771:	gap of 100 bp	
* 41772	42631:	contig of 860 bp	in length
* 42632	42731:	gap of 100 bp	
* 42732	43606:	contig of 875 bp	in length
* 43607	43706:	gap of 100 bp	
* 43707	44539:	contig of 833 bp	in length
* 44540	44639:	gap of 100 bp	
* 44640	45524:	contig of 885 bp	in length
* 45525	45624:	gap of 100 bp	
* 45625	46512:	contig of 888 bp	in length
* 46513	46612:	gap of 100 bp	
* 46613	47479:	contig of 867 bp	in length
* 47480	47579:	gap of 100 bp	
* 47580	48478:	contig of 899 bp	in length
* 48479	48578:	gap of 100 bp	
* 48579	49448:	contig of 870 bp	in length
* 49449	49548:	gap of 100 bp	
* 49549	50430:	contig of 882 bp	in length
* 50431	51430:	gap of 100 bp	
* 51431	51530:	contig of 900 bp	in length
* 51531	52409:	contig of 879 bp	in length
* 52410	52509:	gap of 100 bp	
* 52510	53385:	contig of 876 bp	in length
* 53386	53485:	gap of 100 bp	
* 53486	54376:	contig of 891 bp	in length
* 54377	54476:	gap of 100 bp	
* 54477	55333:	contig of 857 bp	in length
* 55334	55434:	gap of 100 bp	
* 55434	56289:	contig of 856 bp	in length
* 56290	56389:	gap of 100 bp	
* 56390	57282:	contig of 893 bp	in length
* 57283	57382:	gap of 100 bp	
* 57383	58278:	contig of 896 bp	in length
* 58279	58378:	gap of 100 bp	
* 58379	59254:	contig of 876 bp	in length
* 59255	59354:	gap of 100 bp	
* 59355	60220:	contig of 866 bp	in length
* 60221	60320:	gap of 100 bp	
* 60321	61212:	contig of 892 bp	in length
* 61213	61312:	gap of 100 bp	
* 61313	62196:	contig of 884 bp	in length
* 62197	62296:	gap of 100 bp	
* 62297	63167:	contig of 871 bp	in length
* 63168	63267:	gap of 100 bp	
* 63268	64189:	contig of 902 bp	in length
* 64170	64289:	gap of 100 bp	
* 64270	65168:	contig of 899 bp	in length
* 65169	65268:	gap of 100 bp	
* 65269	66137:	contig of 869 bp	in length
* 66138	66237:	gap of 100 bp	
* 66238	67123:	contig of 886 bp	in length
* 67124	67224:	gap of 100 bp	
* 67224	68104:	contig of 881 bp	in length
* 68105	68204:	gap of 100 bp	
* 68205	69094:	contig of 890 bp	in length

Query Match

Best Local Similarity

Matches 297; Conservative

94.9%; Score 284.8; DB 2; Length 73806;

99.0%; Pred. No. 2.9e-78;

0; Mismatches 2; Indels 1; Gaps 1;

QY

1

AAGGAGGAGGAGTGGGCTAGGAGAATCGAGAGATCAGAATTTTAACTCAGCCAGCCA

60

DB

50256

AAGGAGGAGGAGTGGGCTAGGAGAATCGAGAGATCAGAATTTTAACTCAGCCAGCCA

50197

QY

61

TTAATCATGCTCAAGTACTCTCTATCATATTTTGTAGAGACAACAGTTCACCTGAATGAAT

120

DB

50196

TTAATCATGCTCAAGTACTCTCTATCATAT-TGGTAAAGACAACAGTTCACCTGAATGAAT

50138

QY

121

TCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTCTTCTGAGGAAATCTAAGGCAC

180

DB

50137

TCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTCTTCTGAGGAAATCTAAGGCAC

50078

QY

181

AACTGAGGAATGAAGTCAAGGCTTTTCCAAATCCCGAAATACTCTCTCACTGTTACTCATG

240

DB

50077

AACTGAGGAATGAAGTCAAGGCTTTTCCAAATCCCGAAATACTCTCTCACTGTTACTCATG

50018

QY

241

TCCCTTGGAAATTAAGAGGAAGCCAGGAGCATAGTCCCATTAACCAGGAGTGAATCTTCT

300

DB

50017

TCCCTTGGAAATTAAGAGGAAGCCAGGAGCATAGTCCCATTAACCAGGAGTGAATCTTCT

49958

RESULT 8

AX346554

LOCUS

AX346554

Sequence 1625 from Patent WO0200928.

ACCESSION

AX346554

VERSION

AX346554.1

GI:18494440

KEYWORDS

synthetic construct

SOURCE

synthetic construct

ORGANISM

other sequences; artificial sequences.

REFERENCE

1

AUTHORS

Olek, A., Piepenbrock, C. and Berlin, K.

TITLE

Diagnosis of diseases associated with the immune system

JOURNAL

Patent: WO 0200928-A 1625 03-JAN-2002;

FEATURES

Epigenomics AG (DE)

Location/Qualifiers

1..5926

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match

Best Local Similarity

Matches 240; Conservative

68.0%; Score 204; DB 6; Length 5926;

80.0%; Pred. No. 6.9e-53;

0; Mismatches 60; Indels 0; Gaps 0;

QY

1

AAGGAGGAGGAGTGGGCTAGGAGAATCGAGAGATCAGAATTTTAACTCAGCCAGCCA

60

DB

3693

AAGGAGGAGGAGTGGGCTAGGAGAATCGAGAGATCAGAATTTTAACTTATTGTTA

3752

QY

61

TTAATCATGCTCAAGTACTCTCTATCATATTTTGTAGAGACAACAGTTCACCTGAATGAAT

120

DB

3753

TTAATCATGCTCAAGTACTCTCTATCATATTTTGTAGAGACAACAGTTCACCTGAATGAAT

3812

QY

121

TCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTCTTCTGAGGAAATCTAAGGCAC

180

DB

3813

TTTAAAGGTTTTTGGGTTTTTATAGTGTGTTTTTGTAGTTTTTGTAGGAAATTTAAGGTAT

3872

QY

181

AACTGAGGAATGAAGTCAAGGCTTTTCCAAATCCCGAAATACTCTCTCACTGTTACTCATG

240

DB

3873

AACTGAGGAATGAAGTCAAGGCTTTTCCAAATCCCGAAATACTCTCTCACTGTTACTCATG

3932

QY

241

TCCCTTGGAAATTAAGAGGAAGCCAGGAGCATAGTCCCATTAACCAGGAGTGAATCTTCT

300

DB

3933

TTTTTTGGAAATTAAGAGGAAGTGTAGGAAATAGTTGTTATTAATTAGGAGTGAATTTT

3992

```
RESULT 9
AX346555/c
LOCUS AX346555 5926 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1626 from Patent WO0200928.
ACCESSION AX346555
VERSION AX346555.1 GI:18494441
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1
REFERENCE 1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1626 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source Location/Qualifiers
1..5926
/mol_type="synthetic construct"
/organism="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN
Query Match 64.8%; Score 194.4; DB 6; Length 5926;
Best Local Similarity 78.0%; Pred. No. 7.2e-50;
Matches 234; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 1 AGGAGGAGCGTGGCTAGGAAATCGAGATCAGAAATTTAAACTCAGCCGACCA 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2234 AAAAAAAACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2175
QY 61 TTAACATGCTCAAGTACTCCCTATCATATTGTGAGACACACAGTTCAGATGAT 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2174 TTAACATGCTCAAGTACTCCCTATCATATTGTGAGACACACAGTTCAGATGAT 2115
QY 121 TCTAAGGCTCTTTGGGTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2114 TCTAAGGCTCTTTGGGTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 2055
QY 181 AACTGAGGATGAGTCAGGCTTCCAAATCCGAAATATCTCTCCACTGCTTACTCATG 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2054 AACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1995
QY 241 TCCCTTGGAAATTAAGAGGAGCCAGGACATAGCTGCCATAACCCAGGATGAACTTCT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1994 TCCCTTGGAAATTAAGAGGAGCCAGGACATAGCTGCCATAACCCAGGATGAACTTCT 1935
RESULT 10
BOVMCP1X
LOCUS BOVMCP1X 3405 bp DNA linear MAM 17-JAN-1995
DEFINITION Bovine monocyte chemoattractant protein-1 (MCP-1) gene exons 1-3, complete cds.
ACCESSION L32659
VERSION L32659.1 GI:624393
KEYWORDS monocyte chemoattractant protein-1.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1
AUTHORS Wempe,F., Kuhlmann,J.K. and Scheit,K.H.
TITLE Characterization of the bovine monocyte chemoattractant protein-1 gene
JOURNAL Biochem. Biophys. Res. Commun. 202 (3), 1272-1279 (1994)
MEDLINE 94338337
PUBMED 8060303
COMMENT Original source text: Bos taurus male DNA.
FEATURES
source Location/Qualifiers
1..3405
/organism="Bos taurus"
```

```

/mol_type="genomic DNA"
/db_xref="taxon:9913"
/sex="male"
/cell_type="monocyte"
1310..1315
/gene="MCP-1"
/note="putative"
1342..1478
/gene="MCP-1"
/number=1
Join(1403..1478,2227..2344,2695..2800)
/gene="MCP-1"
Join(1403..1478,2227..2344,2695..2800)
/gene="MCP-1"
/codon_start=1
/product="monocyte chemoattractant protein-1"
/protein_id="AA60956.1"
/db_xref="GI:624394"
/translation="MKVSAALLCLLLTVAAFSTFLVLAQPDAINSOVACCYTFNSKKIS
MORLMNRYRVTSSKCPKEAVIFKTLIGKELCADPKQKWQDSINYLKKNQTPKP"
1479..2226
/gene="MCP-1"
/number=1
2227..2344
/gene="MCP-1"
/number=2
2345..2694
/gene="MCP-1"
/number=2
2695..3227
/gene="MCP-1"
/number=3
3202..3207
/gene="MCP-1"
TATA_signal
exon
gene
CDS
intron
exon
intron
exon
polyA_signal
ORIGIN
Query Match 29.2%; Score 87.6; DB 4; Length 3405;
Best Local Similarity 67.9%; Pred. No. 2.7e-16;
Matches 152; Conservative 0; Mismatches 69; Indels 3; Gaps 2;
QY 77 ACTCTATCATATTTGTAGAGACACACAGTTCACCTGAAATGAATTTCTAAGGCTTTTGGGT 136
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 AATTCGTATATTTGTGAAAGAAAGAGTTCATGAAATGAACACTAAGGCTTTTAGG 61
QY 137 TTTTATCATGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCACACAGTGAAGTCACT 196
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 TGTATACGGA--TTCTGGGATTTCTAGGAAAGTCTAAGGCGGAGTTGCAGCATGAAG 119
QY 197 CAGGCTTTCCCAATTCGCCAAATACCTCTCCACTGCTTACTCATGTCCCTTGGAAATTAAG 256
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
120 CGGCCATTCTAATTCAGGAGACGTGTTCTTTGGCTCAGCCCTTCTCTTGGAA-TCCT 178
QY 257 AAGGAAGCCAGAGCATAGCTGCCATAACCCAGGAGTGAACCTTCT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 AAAGAAACTGGGATAACAGTTGCCATAGCCGAGGACACGCTCT 222
RESULT 11
MMU12470
LOCUS MMU12470 2788 bp DNA linear ROD 13-AUG-1994
DEFINITION Mus musculus Balb/c macrophage chemoattractant protein-1 (mcp-1) gene, 5' flanking region.
ACCESSION U12470
VERSION U12470.1 GI:529692
KEYWORDS chemokine; mcp-1; macrophage chemoattractant protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Alberta,J.A., Irminger,J. and Stiles,C.D.
TITLE Macrophage chemoattractant protein-1 (mcp-1) 5' flanking region
JOURNAL Unpublished
```

```

REFERENCE
AUTHORS      2 (bases 2321 to 2787)
TITLE        Rollins,B.J., Morrison,B.D. and Stiles,C.D.
JOURNAL      Cloning and expression of JE, a gene inducible by platelet-derived
MEDLINE      growth factor and whose product has cytokine-like properties
PUBMED       Proc. Natl. Acad. Sci. U.S.A. 85 (11), 3738-3742 (1988)
AUTHORS      3287374
REFERENCE    2 (bases 1 to 2788)
AUTHORS      Alberta,J.A.
TITLE        Direct Submission
JOURNAL      Submitted (20-JUL-1994) John A. Alberta, Cellular and Molecular
JOURNAL      Biology, Dana Farber Cancer Institute and Harvard Medical School,
JOURNAL      44 Binney St., Boston, MA 02115, USA
FEATURES
source       Location/Qualifiers
1..2788
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/clone="JEJCI-50"
/cell_line="Balb/c 3T3 A31"
/cell_type="fibroblast"
2788
/gene="mcp-1"
/genes="mcp-1"
2788..>2788
/gene="mcp-1"
/product="macrophage chemoattractant protein-1"

ORIGIN
Query Match      22.7%; Score 68; DB 10; Length 2788;
Best Local Similarity 59.7%; Pred. No. 4e-10;
Matches 135; Conservative 0; Mismatches 85; Indels 6; Gaps 1;

QY      1  AGGAGGAGCGAGTGGCTAGGAGATCGAGAGATCAGAGATTTTAACTCAGCCAGCCA 60
DB      1609  AAGGATGCTACAGAGGAGGAGGAATCAAGATACCTGAGTGAAGACTCCGCTCAGCCA 1668

QY      61  TTAACATGCTCAAGTACTCTTATCATATTGTGAAGAGACAACAGTTCTCAATGAATGAAT 120
DB      1669  CTGGCAGCATGTGAGAGCGCACTCTTTTAT-----CAAGTCTGCTGAATGAAT 1722

QY      121  TCTAAGCTCTTGGGTTTTATCAGTGTGCTCTCTGTGAGTTCTGAGGAATCTAAGGCAC 180
DB      1723  TCTAAGGCTTTTCAGATTTTATCGCTTTCATCACACTGTTTGTGAAGAAATCTAAACCTG 1782

QY      181  AACTGAGGATGAGTCAGGCTTCCAAATCCGAAATACTCTCTC 226
DB      1783  GAAAGCTGAATTAAGCCAGACATTCAGTTGGCTCACTCATGCTCC 1828

RESULT 12
AC022299/c
LOCUS      AC022299      160869 bp      DNA      linear      HTG 16-OCT-2001
DEFINITION Mus musculus clone RP23-433D8, *** SEQUENCING IN PROGRESS ***, 32
unordered pieces.
ACCESSION  AC022299
VERSION     AC022299, 9 GI:16118057
KEYWORDS    HTG; HTGS PHASE1.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 160869)
AUTHORS    Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
Dederich,D., Thomas,S., Williams,G., Carlock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,J., Liu,W., Loulseghe,H., Lozado,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,

```

```

Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Oguh,M., Parish,B.,
Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstein,G.,
Worley,K. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 160869)
Worley,K.C.
Direct Submission
Submitted (29-JAN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Oct 14, 2001 this sequence version replaced gi:11079355.
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: MANU
Center clone name: RP23-433D8
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye: 20% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143379 bases at least Q40
Consensus quality: 158303 bases at least Q30
Consensus quality: 164427 bases at least Q20
Estimated insert size: 165326; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 32 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*
* 12544: contig of 12544 bp in length
* 12545 12644: gap of unknown length
* 12645 12644: gap of unknown length
* 12645 22802: contig of 10158 bp in length
* 22803 22902: gap of unknown length
* 22903 31226: contig of 8324 bp in length
* 31227 31326: gap of unknown length
* 31327 39328: contig of 8002 bp in length
* 39329 39428: gap of unknown length
* 39429 47713: contig of 8285 bp in length
* 47714 47813: gap of unknown length
* 47814 55305: contig of 7492 bp in length
* 55306 55405: gap of unknown length
* 55406 62668: contig of 7263 bp in length
* 62669 62768: gap of unknown length
* 62769 68691: contig of 5923 bp in length
* 68692 68791: gap of unknown length
* 68792 72517: contig of 3726 bp in length
* 72518 72617: gap of unknown length
* 72618 77895: contig of 5278 bp in length
* 77896 77995: gap of unknown length
* 77996 81790: contig of 3794 bp in length
* 81790 81889: gap of unknown length
* 81890 87749: contig of 5860 bp in length
* 87750 87849: gap of unknown length
* 87850 93175: contig of 5326 bp in length
* 93176 93275: gap of unknown length
* 93276 97924: contig of 4649 bp in length
* 97925 98024: gap of unknown length
* 98025 102235: contig of 4211 bp in length

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

\* 102236' 102335: gap of unknown length  
\* 102336 105916: contig of 3581 bp in length  
\* 105917 106016: gap of unknown length  
\* 106017 110516: contig of 4500 bp in length  
\* 110517 110616: gap of unknown length  
\* 110617 114450: contig of 3834 bp in length  
\* 114451 114550: gap of unknown length  
\* 114551 117745: contig of 3195 bp in length  
\* 117746 117845: gap of unknown length  
\* 117846 121794: contig of 3948 bp in length  
\* 121794 121894: gap of unknown length  
\* 121894 127203: contig of 5310 bp in length  
\* 127203 127303: gap of unknown length  
\* 127304 131479: contig of 4176 bp in length  
\* 131480 131579: gap of unknown length  
\* 131580 135202: contig of 3623 bp in length  
\* 135203 135302: gap of unknown length  
\* 135303 138866: contig of 3564 bp in length  
\* 138867 138966: gap of unknown length  
\* 138967 142626: contig of 3660 bp in length  
\* 142627 142726: gap of unknown length  
\* 142727 146422: contig of 3696 bp in length  
\* 146423 146522: gap of unknown length  
\* 146523 149367: contig of 2745 bp in length  
\* 149368 149367: gap of unknown length  
\* 149368 151602: contig of 2235 bp in length  
\* 151603 151702: gap of unknown length  
\* 151703 153749: contig of 2047 bp in length  
\* 153750 153849: gap of unknown length  
\* 153850 156210: contig of 2361 bp in length  
\* 156211 156310: gap of unknown length  
\* 156311 158696: contig of 2386 bp in length  
\* 158697 158796: gap of unknown length  
\* 158797 160869: contig of 2073 bp in length.

Location/Qualifiers

FEATURES

source

1. 160869  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-433D8"

ORIGIN

Query Match 22.7%; Score 68; DB 2; Length 160869;  
Best Local Similarity 59.7%; Pred. No. 4.5e-10;  
Matches 135; Conservative 0; Mismatches 85; Indels 6; Gaps 1;  
  
QY 1 AAGGAGGCGAGTGGCTAGGAGATCGAGAGATCAGAAATTTTAACTCAGCCAGCCA 60  
DB 122840 AAGGATGCTACAGAGGAGGAGAAATCAAGATACCTGAGTGGAGAACTCCGCTCAGCCA 122781  
  
QY 61 TTAACATGCTCAAGTACTCTATCATATTTCTAAGAGACACAGTTCACTGAATGAAT 120  
DB 122780 CTGGCAGCATGTGAGAGCGGCACATCTTTTAT-----CAAGATCTGCTGAATGAAT 122727  
  
QY 121 TCTAAGGCTTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAATCTAAGGCAC 180  
DB 122726 TCTAAGGCTTTTACAGATTTTATCGCTTTGATCACACTGTTTGTGAAGAAATCTAACCTG 122667  
  
QY 181 AACTGAGGATGAAGTCAGGCTTTCAATTCGGAATACTCCTCC 226  
DB 122666 GAAAGCTGAATTAAGCCAGACATTCAGTTGGCTCACTCATGTGTC 122621

RESULT 13

AC073824 AC073824 180944 bp DNA linear HTG 29-JUN-2000  
DEFINITION Mus musculus clone RP23-92G22, WORKING DRAFT SEQUENCE, 48 unordered  
pieces.  
ACCESSION AC073824  
VERSION AC073824.1 GI:8810441  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
DOE Joint Genome Institute.  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 180944)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov

Project Information  
Center Project Name: 1772366  
Center clone name: RPCI-23\_92G22

Summary Statistics

Consensus quality: 143655 bases at least Q40  
Consensus quality: 160562 bases at least Q30  
Consensus quality: 164105 bases at least Q20  
Estimated insert size: 194000; pulse field gel estimation  
Estimated insert size: 176244; sum-of-contigs estimation  
Quality coverage: 6.92 in Q20 bases; pulse field gel estimation  
Quality coverage: 7.62 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 48 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1102: contig of 1102 bp in length  
\* 1103 1202: gap of unknown length  
\* 1203 2608: contig of 1406 bp in length  
\* 2609 2708: gap of unknown length  
\* 2709 3827: contig of 1119 bp in length  
\* 3828 3927: gap of unknown length  
\* 3928 5266: contig of 1339 bp in length  
\* 5267 5366: gap of unknown length  
\* 5367 6856: contig of 1490 bp in length  
\* 6857 6956: gap of unknown length  
\* 6957 8006: contig of 1050 bp in length  
\* 8007 8106: gap of unknown length  
\* 8107 9357: contig of 1251 bp in length  
\* 9358 9457: gap of unknown length  
\* 9458 10998: contig of 1541 bp in length  
\* 10999 11098: gap of unknown length  
\* 11099 12370: contig of 1272 bp in length  
\* 12370 12470: gap of unknown length  
\* 12471 13567: contig of 1097 bp in length  
\* 13568 13667: gap of unknown length  
\* 13668 15007: contig of 1340 bp in length  
\* 15008 15107: gap of unknown length  
\* 15108 17117: contig of 2010 bp in length  
\* 17118 17217: gap of unknown length  
\* 17218 18346: contig of 1129 bp in length  
\* 18347 18446: gap of unknown length  
\* 18447 19824: contig of 1378 bp in length  
\* 19825 19924: gap of unknown length  
\* 19925 21617: contig of 1693 bp in length  
\* 21618 21717: gap of unknown length  
\* 21718 23913: contig of 2196 bp in length  
\* 23914 24013: gap of unknown length  
\* 24014 25142: contig of 1129 bp in length  
\* 25143 25242: gap of unknown length  
\* 25243 26540: contig of 1298 bp in length  
\* 26541 26640: gap of unknown length  
\* 26641 28457: contig of 1817 bp in length





Tue Aug 9 17:01:27 2005

Best Local Similarity	59.7%;	Pred. No. 4.6e-10;					
Matches 195;	Conservative	0;	Mismatches	85;	Indels	6;	Gaps 1;
1	AAGGAGGAGGCACGTGGGCTAGGAGAAATCAGAGATCAGAAATTTAAATCTAGCCAGCCCA	60					
103778	AAGGAGTCTACAGAGGAAGGAGAAATCAAGATACCTGAGTGGAGAGATCTCGCTCAGCCA	103719					
61	TTAAACATGCCCTCAAGTACTCTCTATCATATTTTGAAGAGACAACAGTTCACTGAATGAAT	120					
103718	CTGGCAGCAGATGTGAGAGGCCACTCTTTTAT-----CAAGAGTCTGCTGAATGAAT	103665					
121	TCTAAGTCTCTTGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAATCTAAGGCAC	180					
103664	TCTAAGGCCTTCAGATTTTATCGCTTTGATCACATGTGTTGTGAAGAANTCTAAACCTG	103605					
181	AACCTGAGGAATGAGTCAGGCTTTTCCAATTCGCGAAATACTCTCTC	226					

RESULT 15	223726 bp	DNA	linear	ROD 05-APR-2002
AL626807/c				
LOCUS				
DEFINITION	Mus DNA sequence from clone Rp23-350g1 on chromosome 11, complete sequence.			
ACCESSION	AL626807			
VERSION	AL626807.7			
KEYWORDS	GI:17221258			
SOURCE	Mus musculus (house mouse)			
ORGANISM				
	Mus musculus			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1			
REFERENCE	Whitehead,S.			
AUTHORS	Direct Submission			
TITLE	Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,			
JOURNAL	Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk			
	humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
	On Nov 30, 2001 this sequence version replaced gi:17065774.			
COMMENT	from overlapping clones.			

variation annotation may not be found in the sequence subassembly corresponding to the overlapping clone, as we submit subassemblies only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Swi: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-350G1 is from the RP23-23 Mouse PAC Library constructed by the group of Pieter de Jong.

```

FEATURES
  source
    Location/Qualifiers
      1. 223726
        /organism="Mus musculus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10090"
        /chromosome="11"
        /clone="RP23-350G1"
        /clone_lib="RPCI-23"

ORIGIN
  Query Match      22.7% ; Score 68 ; DB 10 ; Length 223726 ;
  Best Local Similarity 59.7% ; Pred. No. 4.6e-10 ;

```

Matches	135;	Conservative	0;	Mismatches	85;	Indels	6;	Gaps	1;
1	AAGGAGGAGCAGCTGGCTAGGAGAA	TCCGAGAGATCAGAAATTTTAA	ACTCAGGCCGACCA	60					
12956	AAGGATGCTTCAGAGGAGAGGAGNAATCAAGATCCTGTGCTGGAAGACTCCGCTCAGCCA	12897							
61	TTAAATCATGCCCTCAAGTACTCCTATCATATTTGTGAAGAGACAAACAGATTCTCAATGAAT	120							
12896	CTGGCAGCGCATGTGAGAGCGCCACTCTTTTAT-----CAAGAGTCTGCTGAAATGAAT	12843							
121	TCTAAGTCTTTGGTGTATTTATCAGTGTGCTTCTGTATGTTTCTGAGAAATCTTAAGGCAC	180							
12842	TCTAAGGCTTTTCAGATTTTATCGCTTTGATCACA	CTGTTGTGAAGAAATCTTAACCTG	12783						
181	AACTGAGAAATGAAGTCAAGCTTCAGGCTTTCCAAATTC	CCCGAAATAC	TCCTCC	226					
12782	GAAAGCTGAATTAAGCCAGACATTCAGTTGGCTCACTCATGGTCC	12737							

Search completed: August 4, 2005, 11:16:54  
Job time : 1209.83 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 198.826 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276c-5  
Perfect score: 300  
Sequence: 1 aaggaggaggagtgaggcta.....ataaccaggatgaactct 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	ADH13942	Adh13942 Human mon
2	296.8	98.9	3221	12 ADN12120	Adn12120 MCP1 gene
3	296.8	98.9	3221	12 ADO03802	Ado03802 Human Ccl1
4	296.8	98.9	11793	12 ADO03803	Ado03803 Human Ccl1
5	204	68.0	5926	6 ABL33852	Ab133852 Human imm
6	194.4	64.8	5926	6 ABL33853	Ab133853 Human imm
7	36.8	12.3	204621	11 ACN44486	Acn44486 Human gen
8	35.4	11.8	2122	6 ABK52382	Abk52382 DNA encod
9	35	11.7	32186	5 AAS374422	Aas374422 Human DNA
10	35	11.7	38855	12 ADN41808	Adn41808 Novel hum
11	34.4	11.5	110000	6 ABX08336	Continuation (8 of
12	34.4	11.5	110000	12 ADJ25985_07	Continuation (8 of
13	34.4	11.5	110000	12 ADN97989_07	Continuation (8 of
14	34.4	11.5	110000	12 ADO50281_07	Continuation (8 of
15	33.8	11.3	2268	5 AAS70462	Aas70462 DNA encod
16	33.8	11.3	2268	5 AAS72056	Aas72056 DNA encod
17	33.6	11.2	693	8 ACN30296	Acn30296 Prokaryot
18	33.2	11.1	34548	11 ACN44070	Acn44070 Human gen
19	32.8	10.9	17918	6 AAS61419	Aas61419 Human gen
20	32.8	10.9	36485	4 AAK74751	Aak74751 Human imm

C	21	32.8	10.9	36485	4	AAK68958	Aak68958 Human imm
	22	32.6	10.9	661	12	ADK72196	Adk72196 Antimicro
	23	32.6	10.9	661	12	ADK72243	Adk72243 Antimicro
	24	32.6	10.9	3196	4	AAL04311	Aal04311 Human rep
C	25	32.6	10.9	4866	2	AAQ82848	Aaq82848 P-selecti
C	26	32.6	10.9	4866	3	AAA34926	Aaa34926 Human ade
C	27	32.6	10.9	4866	3	AAF21048	Aaf21048 Human low
C	28	32.6	10.9	4866	10	ABZ96742	Abz96742 Human nuc
C	29	32.6	10.9	4866	11	ABD20591	Abd20591 Human pul
	30	32.6	10.9	18595	4	AAS33411	Aas33411 DNA encod
C	31	32.6	10.9	78025	8	ABQ77404	Abq77404 Human SEL
	32	32.6	10.9	105219	11	ACN44286	Acn44286 Human gen
	33	32.6	10.9	110469	12	ADQ97337	Adq97337 Human can
	34	32.6	10.9	262090	12	ADO59207	Ado59207 MSI-H car
C	35	32.4	10.8	393	3	AAH30439	Aah30439 Human col
C	36	32.4	10.8	10159	4	AAK73470	Aak73470 Human imm
C	37	32.4	10.8	10159	4	AAK73471	Aak73471 Human imm
C	38	32.4	10.8	24458	12	ADQ97473	Adq97473 Human can
	39	32.4	10.8	84073	11	ACN44322	Acn44322 Human gen
	40	32.4	10.8	149612	11	ACN45154	Acn45154 Human gen
C	41	32.2	10.7	5483	12	ADG47889	Adg47889 Arabidops
C	42	32	10.7	346	10	ADK61209	Adk61209 Ovarian c
	43	32	10.7	496	4	AAH32965	Aah32965 Human col
C	44	32	10.7	4199	4	ABL17530	Ab117530 Drosophil
	45	32	10.7	37590	4	AAS12439	Aas12439 DNA encod

## ALIGNMENTS

## RESULT 1

ADH13942  
ID ADH13942 standard; DNA; 300 BP.  
XX  
AC ADH13942;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:5.  
XX  
KW ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
KW cervical carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN EP1170372-Al.  
XX  
PD 09-JAN-2002.  
XX  
PF 06-JUL-2000; 2000EP-00114560.  
XX  
PR 06-JUL-2000; 2000EP-00114560.  
XX  
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
PI Roessel F, Soto U, Coy J, Finzer P, Delius H, Poustka A;  
PI Zur Hausen H, Patzelt A;  
XX  
WPI; 2002-165895/22.  
XX  
PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
XX useful in gene therapy, for treating atherosclerosis and cancer.  
XX  
PS Disclosure; SEQ ID NO 5; 30pp; English.  
XX  
CC The invention relates to a novel nucleic acid molecule (I) comprising a  
CC sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
CC having the biological activity of MCP-1. A protein encoded by a nucleic  
CC acid of the invention has cytostatic, and antiarteriosclerotic activity.  
CC A nucleic acid of the invention may have a use in gene therapy. A  
CC compound of the invention is useful in the preparation of a medicament  
CC for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

Tue Aug 9 17:01:27 2005

CC pharmacological composition of the invention is useful for the treatment  
CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
CC atherosclerosis or cancer. The present sequence is used in the  
CC exemplification of the invention.

XX  
SQ Sequence 300 BP; 92 A; 62 C; 67 G; 79 T; 0 U; 0 Other;  
Query Match 100.0%; Score 300; DB 6; Length 300;  
Best Local Similarity 100.0%; Pred. No. 4.6e-86;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 60  
DB 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 60

QY 61 TTAACATGCTCAAGTACTCCCTATCATATTTGTAAGAGACAACAGTTCACTGAAATGAAT 120  
DB 61 TTAACATGCTCAAGTACTCCCTATCATATTTGTAAGAGACAACAGTTCACTGAAATGAAT 120

QY 121 TCTAAGGTCTTTGGGTTTTTATCAGTGTCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 180  
DB 121 TCTAAGGTCTTTGGGTTTTTATCAGTGTCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 180

QY 181 AACTGAGGAGTGAAGTCAGGCTTTCCAAATCCCGAAATATCTCCTCCACTGCTTACTCATG 240  
DB 181 AACTGAGGAGTGAAGTCAGGCTTTCCAAATCCCGAAATATCTCCTCCACTGCTTACTCATG 240

QY 241 TCCCTTTGGAAATTAAGAGAGCCAGGAGCATAGTGCATACCAACAGGATGAATCTTCT 300  
DB 241 TCCCTTTGGAAATTAAGAGAGCCAGGAGCATAGTGCATACCAACAGGATGAATCTTCT 300

RESULT 2  
ADN12120  
ID ADN12120 standard; DNA; 3221 BP.  
XX  
AC ADN12120;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE MCP1 gene promoter region.  
XX  
KW major histocompatibility class I; MHC-I; MHC-II; Cytostatic;  
KW EBV-associated cancer; Hodgkin's lymphoma; nasopharyngeal carcinoma;  
KW gastric carcinoma; Burkitt's lymphoma; T-cell lymphoma; B-cell lymphoma;  
KW parotid carcinoma; breast carcinoma; leiomyosarcoma; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2004027036-A2.  
XX  
PD 01-APR-2004.  
XX  
PF 19-SEP-2003; 2003WO-US029684.  
XX  
PR 19-SEP-2002; 2002US-0411990P.  
XX  
PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
XX  
PI Ambinder RF, Yang Y, Borrello IM, Levitsky HI;  
XX  
DR WPI; 2004-295406/27.  
XX  
XX New human cell line modified to comprise and express genes encoding  
PT immunomodulators and an antigen of Epstein-Barr virus (EBV), useful for  
PT inducing or stimulating an immune response in a human to EBV-associated  
PT cancer.  
XX  
XX Disclosure; SEQ ID NO 13; 218pp; English.  
PS  
XX The present invention relates to a human cell line, which lacks major  
CC histocompatibility class I (MHC-I) and MHC-II antigens and which has been  
CC modified to comprise and express a gene encoding an immunomodulator and a

CC gene encoding an antigen of Epstein-Barr virus (EBV). The human cell  
CC line, compositions and methods are useful for inducing or stimulating an  
CC immune response in a human to an EBV-associated cancer, where the human  
CC has or is at risk for Hodgkin's lymphoma, nasopharyngeal carcinoma,  
CC gastric carcinoma, Burkitt's lymphoma, T-cell lymphoma, B-cell lymphoma,  
CC parotid carcinoma, breast carcinoma, and leiomyosarcoma. The present  
CC sequence represents a nucleotide sequence associated with the cell line  
CC of the invention.

XX  
SQ Sequence 3221 BP; 859 A; 793 C; 727 G; 842 T; 0 U; 0 Other;  
Query Match 98.9%; Score 296.8; DB 12; Length 3221;  
Best Local Similarity 99.3%; Pred. No. 1.3e-84;  
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 60  
DB 1445 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCA 1504

QY 61 TTAACATGCTCAAGTACTCCCTATCATATTTGTAAGAGACAACAGTTCACTGAAATGAAT 120  
DB 1505 TTAACATGCTCAAGTACTCCCTATCATATTTGTAAGAGACAACAGTTCACTGAAATGAAT 1564

QY 121 TCTAAGGTCTTTGGGTTTTTATCAGTGTCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 180  
DB 1565 TCTAAGGTCTTTGGGTTTTTATCAGTGTCTTCTGAGTCTTCTGAGGAAATCTAAGGCAC 1624

QY 181 AACTGAGGAGTGAAGTCAGGCTTTCCAAATCCCGAAATATCTCCTCCACTGCTTACTCATG 240  
DB 1625 AACTGAGGAGTGAAGTCAGGCTTTCCAAATCCCGAAATATCTCCTCCACTGCTTACTCATG 1684

QY 241 TCCCTTTGGAAATTAAGAGAGCCAGGAGCATAGTGCATACCAACAGGATGAATCTTCT 300  
DB 1685 TCCCTTTGGAAATTAAGAGAGCCAGGAGCATAGTGCATACCAACAGGATGAATCTTCT 1744

RESULT 3  
ADO03802  
ID ADO03802 standard; DNA; 3221 BP.  
XX  
AC ADO03802;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human Cc12 promoter region used for gene knockout animal models SeqID 3.  
XX  
KW human; ds; animal model; age-related macular degeneration; AMD;  
KW gene knockout; Cc12-deficient; Cor2-deficient; drusen;  
KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;  
KW choroidal neovascularisation; ophthalmological; gene therapy; promoter.  
XX  
OS Homo sapiens.  
XX  
PN WO2004041160-A2.  
XX  
PD 21-MAY-2004.  
XX  
PF 16-OCT-2003; 2003WO-US032933.  
XX  
PR 30-OCT-2002; 2002US-0422096P.  
XX  
PA (KENT ) UNIV KENTUCKY RES FOUND.  
XX  
PI Ambati J;  
XX  
DR WPI; 2004-400512/37.  
XX  
PT Testing candidate drug for treating age-related macular degeneration, by  
PT administering drug to Cc12-deficient, Cor2-deficient knockout mouse, and  
PT analyzing development or regression of drusen and/or lipofuscin  
PT accumulation in eye.  
XX  
XX Disclosure; SEQ ID NO 3; 64pp; English.  
PS



```

PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 1625; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 5926 BP; 1562 A; 55 C; 1318 G; 2991 T; 0 U; 0 Other;
XX
Query Match 68.0%; Score 204; DB 6; Length 5926;
Best Local Similarity 80.0%; Pred. No. 1.2e-54;
Matches 240; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 1 AAGGAGGAGCGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAACTCAGCCGACCA 60
DB 3693 AAGGAGGAGTGTGGGTAGGAGATCGAGAGATCAGAAATTTAAATTTAGTTTGA 3752
QY 61 TTAACATGCTCAAGTACTCCTATCATATTGTGAAGAGACAACAGTTCACTGAAATGAAT 120
DB 3753 TTAATATGTTTAAAGTATTTTATATATATTTGTAGAGATATAGTTATTCGAATGAAT 3812
QY 121 TCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180
DB 3813 TTTAAGGTTTTGGGTTTTTATTAGTGTGTTTTTGTAGTTTGTGAGGAAATTTAAGGTAT 3872
QY 181 AACTGAGGAATGAAGTCAGGCTTCCAAATCCCGAATATCTCTCCACTGCTTACTCATG 240
DB 3873 AATTGAGGAATGAAGTTAGGTTTTTAAATTTTCGAAATATTTTATTGTTTATTATG 3932
QY 241 TCCCTTGGAATTAAGAAAGGAGCCGAGGAGCATAGCTGCCATACCCAGGAGTGAATCTT 300
DB 3933 TTTTTCGAAATTAAGAGGAGGAGTTAGGAGATAGTTGTTATATTAGGAGTGAATTTT 3992
RESULT 6
ABL33653/c
ID ABL33653 standard; DNA; 5926 BP.
XX
XX ABL33653;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 1626.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; anti-anaemic; cytostatic; nootropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antididiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX PN

```

```

XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 1626; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 5926 BP; 1554 A; 55 C; 1492 G; 2825 T; 0 U; 0 Other;
XX
Query Match 64.8%; Score 194.4; DB 6; Length 5926;
Best Local Similarity 78.0%; Pred. No. 1.4e-51;
Matches 234; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAACTCAGCCGACCA 60
DB 2234 AAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2175
QY 61 TTAACATGCTCAAGTACTCCTATCATATTGTGAAGAGACAACAGTTCACTGAAATGAAT 120
DB 2174 TTAACATACCTCAATATCTCTATCATATTATATAAACAACAATTCACATAAATAAAT 2115
QY 121 TCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180
DB 2114 TCTAATAATCTTTAAATTTTATCAATATACCTTCTATAATTTCTAAAAAATACTAAAAAC 2055
QY 181 AACTGAGGAATGAAGTCAGGCTTTCCTCAATCCCGAATATCTCTCCACTGCTTACTCATG 240
DB 2054 AACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1995
QY 241 TCCCTTGGAATTAAGAAAGGAGCCGAGGAGCATAGCTGCCATACCCAGGAGTGAATCTT 300
DB 1994 TCCCTTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1935
RESULT 7
ACN44486/c
ID ACN44486 standard; DNA; 204621 BP.
XX
XX ACN44486;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human genomic sequence hCG40063.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
XX PD

```

XX 28-FEB-2003; 2003WO-US006235.  
 XX 01-MAR-2002; 2002US-00087192.  
 XX (SAGR-) SAGRES DISCOVERY.  
 XX Morris DW;  
 XX WPI; 2003-328604/31.  
 XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
 XX comprises a nucleotide sequence.  
 XX Claim 1; SEQ ID NO 958; Opp; English.  
 XX The present invention relates to novel DNA and protein sequences which  
 XX are associated with carcinomas. The sequences are useful for: (i) for  
 XX screening drug candidates; (ii) for screening of bioactive agent capable  
 XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
 XX a bioactive agent capable of modulating the activity of CAP; (iv) for  
 XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
 XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
 XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
 XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
 XX determining Carcinoma Associated (CA) gene copy number. In addition, the  
 XX CA genes are useful as DNA vaccines and the CAP are useful as markers of  
 XX carcinoma including lymphoma. The present sequence is one such CA coding  
 XX sequence. Note: This patent is an equivalent to basic patent  
 XX US2002182586A1, for which no sequence data was published  
 XX SQ Sequence 204621 BP; 53683 A; 43334 C; 43489 G; 57468 T; 0 U; 6647 Other;  
 Query Match 12.3%; Score 36.8; DB 11; Length 204621;  
 Best Local Similarity 58.0%; Pred. No. 2.9;  
 Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
 QY 42 TTAACTCAGCCAGCCATTAACTGCTCAAGTCTCTCTATCATATTTGTAAGAGACA 101  
 Db 105675 TTCAGATCACACAGCATGAAGTTATGATTCCTCTTTTATGCTTTGAAGACATA 105616  
 QY 102 ACAGTTCACTGAATGAATCTTAAGTCTTTGGTTTTTATCAGTGTGCTTC 153  
 Db 105615 ATACTTATGAGACCAATTCAGAGTCATGCGATTTTCAGCAGGTGAGTC 105564  
 RESULT 8  
 ID ABK52382 standard; DNA; 2122 BP.  
 AC ABK52382;  
 XX 27-AUG-2002 (first entry)  
 XX DNA encoding fucose-specific lectin protein.  
 XX Fucose-specific lectin protein; enzyme; foodstuff; drug; cosmetic; gene;  
 XX ds.  
 XX Aspergillus oryzae.  
 XX Key Location/Qualifiers  
 XX CDS 463..1613  
 XX /tag= a  
 XX /product= "fucose lectin protein"  
 XX exon 463..498  
 XX /tag= b  
 XX /number= 1  
 XX /tag= c  
 XX intron 499..549  
 XX /tag= c  
 XX /number= 1  
 XX exon 550..726  
 XX /tag= b

FT intron /number= 2  
 FT 727..776  
 FT /tag= c  
 FT /number= 2  
 FT 777..992  
 FT /tag= b  
 FT /number= 3  
 FT 993..1046  
 FT /tag= c  
 FT /number= 3  
 FT 1047..1148  
 FT /tag= b  
 FT /number= 4  
 FT 1149..1211  
 FT /tag= c  
 FT /number= 4  
 FT 1212..1613  
 FT /tag= b  
 FT /number= 5  
 JP2002112786-A.  
 PD 16-APR-2002.  
 PF 06-OCT-2000; 2000JP-00307979.  
 PR 06-OCT-2000; 2000JP-00307979.  
 XX (GEKK-) GEKKEIKAN KK.  
 XX WPI; 2002-458866/49.  
 DR P-PSDB; AAU97833.  
 XX Fucose-specific lectin protein with enzymic activity useful for  
 XX foodstuffs, drugs and in cosmetics.  
 XX Claim 2; Fig 3-4; 14pp; Japanese.  
 XX The invention describes a fucose-specific lectin protein with enzymic  
 XX activity useful for foodstuffs, drugs and cosmetics. This sequence  
 XX encodes a fucose-specific lectin protein described in the invention  
 XX SQ Sequence 2122 BP; 575 A; 484 C; 505 G; 558 T; 0 U; 0 Other;  
 Query Match 11.8%; Score 35.4; DB 6; Length 2122;  
 Best Local Similarity 54.1%; Pred. No. 1.3;  
 Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
 QY 118 AATCTAAGGTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTTGAGGAATCTAAGG 177  
 Db 1079 AACTCCAAGTTTGTGCTCTCTTCCAGCGCTCTCTGTGCGGATTTAGGAACCATCA 1020  
 QY 178 CACAATCTGAGGAATGAAGTCAAGGCTTTCGAATTCGGAATCTCCTCCACTGCTTACTC 237  
 Db 1019 CCGAAATTGGGAATTTTCAAGTGCCTTACCATTTCCACATATACTCTGTATGTTATC 960  
 QY 238 ATGTCCTTGGAA 250  
 Db 959 AGTTTCTGGGCA 947  
 RESULT 9  
 ID AAS34422/c  
 ID AAS34422 standard; DNA; 32186 BP.  
 XX AAS34422;  
 XX 17-DEC-2001 (first entry)  
 XX Human DNA for a novel foetal antigen, SEQ ID NO 1846.  
 XX Human; foetal tissue antigen; ds; antiinflammatory; neuroprotective;  
 XX immunomodulator; cardiovascular; cytostatic; nephrothropic;

Tue Aug 9 17:01:27 2005

us-09-899-276c-5.rng

KW cardiovascular; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; breast neoplasm; cancer;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; food additive.

XX OS Homo sapiens.

XX PN WO200155312-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001321.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0225113P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 25-SEP-2000; 2000US-0235484P.  
PR 26-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.



us-09-899-276c-5.rng

Tue Aug 9 17:01:27 2005

XX Fischer CB, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z;  
 PI Lafleur DW, Moore PA, Shi Y, Olsen H, Ebner R, Birse CE;  
 XX WPI; 2004-225733/21.  
 DR New isolated nucleic acid encoding human proteins, useful for treating,  
 XX preventing or diagnosing e.g. rheumatoid arthritis, multiple sclerosis,  
 PT anemia, inflammatory bowel disease, atherosclerosis, cancers, chronic  
 PT kidney failure.

XX Disclosure; SEQ ID NO 930; 372pp; English.

XX The invention describes novel human secreted proteins and the nucleotides  
 CC encoding them. The polynucleotides are useful in chromosome  
 CC identification. The polynucleotides are useful in controlling gene  
 CC expression, in gene therapy or as molecular weight markers. The  
 CC polynucleotides and polypeptides are useful for diagnosing, treating or  
 CC preventing diseases of the immune system, immunodeficiencies, e.g.  
 CC Chediak-Higashi syndrome, autoimmune diseases, e.g. systemic lupus  
 CC erythematosus, rheumatoid arthritis, multiple sclerosis, haemolytic  
 CC anaemia or myasthenia gravis, allergic reactions, e.g. asthma,  
 CC inflammatory conditions, e.g. inflammatory bowel disease. They can also  
 CC be used as a stimulator of B cell responsiveness to pathogens or as an  
 CC activator of T cells. The polynucleotides and polypeptides are also  
 CC useful for treating or preventing blood-related disorders, e.g.  
 CC eosinophilia, thrombosis, thromboembolism, atherosclerosis, myocardial  
 CC infarction, unstable angina or anaemia. They can also be used for  
 CC treating, preventing or diagnosing hyperproliferative disorders  
 CC (cancers), renal disorders (chronic kidney failure, renal tubular  
 CC acidosis or kidney stones), cardiovascular disorders or respiratory  
 CC disorders. This sequence represents a novel human secreted protein  
 CC polynucleotide fragment. Note: This sequence is available in electronic  
 CC format from the US patent office at  
 CC ftp.seqdata.uspto.gov/sequence.html?docID=20040044191.

XX Sequence 3885 BP; 11951 A; 6822 C; 6960 G; 13122 T; 0 U; 0 Other;

Query Match 11.7%; Score 35; DB 12; Length 38855;  
 Best Local Similarity 53.2%; Pred. No. 5.7;  
 Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
 QY 70 CTCAGTACTCTATCATTTTGTAGAGACACAGTTCCTCACTGAATGAATTTCTAAGGTC 129  
 Db 13741 CTCATATATACATTTAAATTTTTCACATATATATTTTTCAGTATTTTGCCTTGGATCTTT 13682  
 QY 130 TTGGGCTTTTATCATGTGCTCTGTAGTTTCTGAGAAATCTAAGGCACACTGAGGA 189  
 Db 13681 AATTTTTCATCATCTGCCATATATTTTCTAAGAAATACTGTAAATTTGAATA 13622  
 QY 190 ATGAAGTCAGGCTTTCCAA 208  
 Db 13621 AATAAATGGGTTCCAAA 13603

RESULT 11  
 ABX08336\_07  
 Continuation (8 of 17) of ABX08336 from base 700001 (Human phosphodiesterase 4D (PDE4D))  
 WP Sequence split into 17 fragments LOCUS ABX08336 Accession Abx08336  
 WP Fragment Name Begin End  
 WP ABX08336\_00 1 110000  
 WP ABX08336\_01 100001 210000  
 WP ABX08336\_02 200001 310000  
 WP ABX08336\_03 300001 410000  
 WP ABX08336\_04 400001 510000  
 WP ABX08336\_05 500001 610000  
 WP ABX08336\_06 600001 710000  
 WP ABX08336\_07 700001 810000  
 WP ABX08336\_08 800001 910000  
 WP ABX08336\_09 900001 1010000  
 WP ABX08336\_10 1000001 1110000  
 WP ABX08336\_11 1100001 1210000  
 WP ABX08336\_12 1200001 1310000

WP ABX08336\_13 1300001 1410000  
 WP ABX08336\_14 1400001 1510000  
 WP ABX08336\_15 1500001 1610000  
 WP ABX08336\_16 1600001 1691080  
 Query Match 11.5%; Score 34.4; DB 6; Length 110000;  
 Best Local Similarity 53.8%; Pred. No. 13;  
 Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
 QY 89 TTTCTAGAGACACAGTTCACGAAATGAATCTTAAGTCTTTGGTGTATATCAGTGT 148  
 Db 99904 TTTTGTAGAAATTAATAATTTGAAATTAACCTCTTTTTCAGCCTTTTGAGTTTAT 99963  
 QY 149 GCTTCTGTAGTTCTGAGAAATCTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208  
 Db 99964 GTTCTATAGTTTAGGCTTAAATAATATAGACATTAGAAAAAACTCAAGCGGTATTA 100023  
 QY 209 TTCCCGAAATAC 220  
 Db 100024 ATCCTTATATCC 100035

RESULT 12  
 ADJ25985\_07  
 Continuation (8 of 17) of ADJ25985 from base 700001 (Human phosphodiesterase 4D (PDE4D))  
 WP Sequence split into 17 fragments LOCUS ADJ25985 Accession Adj25985  
 WP Fragment Name Begin End  
 WP ADJ25985\_00 1 110000  
 WP ADJ25985\_01 100001 210000  
 WP ADJ25985\_02 200001 310000  
 WP ADJ25985\_03 300001 410000  
 WP ADJ25985\_04 400001 510000  
 WP ADJ25985\_05 500001 610000  
 WP ADJ25985\_06 600001 710000  
 WP ADJ25985\_07 700001 810000  
 WP ADJ25985\_08 800001 910000  
 WP ADJ25985\_09 900001 1010000  
 WP ADJ25985\_10 1000001 1110000  
 WP ADJ25985\_11 1100001 1210000  
 WP ADJ25985\_12 1300001 1410000  
 WP ADJ25985\_13 1400001 1510000  
 WP ADJ25985\_14 1500001 1610000  
 WP ADJ25985\_15 1600001 1691139  
 WP ADJ25985\_16 1600001 1691139  
 Query Match 11.5%; Score 34.4; DB 12; Length 110000;  
 Best Local Similarity 53.8%; Pred. No. 13;  
 Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 89 TTTGTAGAGACACAGTTCACGAAATGAATCTAAGTCTTTGGTGTATATCAGTGT 148  
 Db 99963 TTTTGTAGAAATTAATAATTTGAAATTAACCTCTTTTTCAGCCTTTTGAGTTTAT 100022  
 QY 149 GCTTCTGTAGTTCTGAGAAATCTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208  
 Db 100023 GTTCTATAGTTTAGGCTTAAATAATATAGACATTAGAAAAAACTCAAGCGGTATTA 100082  
 QY 209 TTCCCGAAATAC 220  
 Db 100083 ATCCTTATATCC 100094

RESULT 13  
 ADN97989\_07  
 Continuation (8 of 17) of ADN97989 from base 700001 (Human phosphodiesterase 4D genome)  
 WP Sequence split into 17 fragments LOCUS ADN97989 Accession Adn97989  
 WP Fragment Name Begin End  
 WP ADN97989\_00 1 110000  
 WP ADN97989\_01 100001 210000  
 WP ADN97989\_02 200001 310000  
 WP ADN97989\_03 300001 410000  
 WP ADN97989\_04 400001 510000  
 WP ADN97989\_05 500001 610000

WP	ADN97989_06	600001	710000
WP	ADN97989_07	700001	810000
WP	ADN97989_08	800001	910000
WP	ADN97989_09	900001	1010000
WP	ADN97989_10	1000001	1110000
WP	ADN97989_11	1100001	1210000
WP	ADN97989_12	1200001	1310000
WP	ADN97989_13	1300001	1410000
WP	ADN97989_14	1400001	1510000
WP	ADN97989_15	1500001	1610000
WP	ADN97989_16	1600001	1691138

  

Query Match	11.5%;	Score 34.4;	DB 12;	Length 110000;
Best Local Similarity	53.8%;	Pred. No. 13;		
Matches	71; Conservative	0; Mismatches	61; Indels	0; Gaps

  

Qy	89	TTTCTGAGAGACAACAGTTCAC	TGAATTAATTCTAAGGTCTTTGGTTTTATCAGTGT	148
Dd	99862	TTTTTTAGAATTTACTAATTTTGAAATAATTAACCTCTCTTTTCAGCCCTTTTGAGTTTTAT	100021	
Qy	149	GCTTCTGTAGTTTCTGAGGAATCTAAGGCACA	CTGAGGAATGAAGTCAGGCTTTCCAA	208
Dd	100022	GTTTCTATAGTTTAGGCTTAABAAATAA	VTAGACATTAGAAAAAACCTCAAGCGGTATTA	100081
Qy	209	TTCCCGAAATAC	220	
Dd	100082	ATCCTTATATCC	100093	

AC	AA570462;
XX	
XX	13-FEB-2002 (first entry)
XX	
XX	DNA encoding novel human diagnostic protein #6266.
XX	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
XX	Homo sapiens.
OS	
XX	WO200175067-A2.
XX	
XX	11-OCT-2001.
PD	
XX	
XX	30-MAR-2001; 2001WO-US0008631.
PF	
XX	
XX	31-MAR-2000; 2000US-00540217.
PR	
PR	23-AUG-2000; 2000US-00649167.
XX	
XX	(HYSE-) HYSEQ INC.
PA	
XX	
FI	Dzmanac RT, Liu C, Tang YT;
XX	
XX	WPI, 2001-639362/73.
DR	P-PSDE; ABG06275.
DR	
XX	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics; forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
XX	Claim 1; SEQ ID NO 6266; 103pp; English.
PS	

Query Match	11.3%	Score	33.8;	DB	5;	Length	2268;
Best Local Similarity	55.6%;	Pred.	No. 4.5;				
Matches	65;	Conservative	0;	Mismatches	52;	Indels	0;
Gaps	0;						
Qy	65	CATGCCCTCAAGTACTCCTATCATATTGTTAAGAGACAACAGTCACGTGAATGAATCTTA	124				
Dd	449	CGTGTTAAATCTCCCATATTATTGTGGGAGTCTAAGTCTCTTTGTAGATCTCTA	390				
Qy	125	AGGTCTTTGGGTTTTATCAGTGTGCCTCTGTAGTCTTCTGAGGAAAATCTAAGGCACA	181				
Dd	389	AGGACTCTCTTTATGAATCTGGGTGCTCTGCTATTGGGTGCATATATTTAGGATA	333				
Search completed: August 4, 2005, 08:31:19							
Job time : 203.826 secs							

us-09-899-276c-5.rng

Tue Aug 9 17:01:27 2005

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 1415.95 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-5  
Perfect score: 300  
Sequence: 1 aaggaggagcgagtggtta.....ataaccaggatgaacttct 300

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.4	14.5	737	9	CR231209 Reverse s
C 2	39	13.0	648	8	AZ446937 IM0243N04
C 3	37.8	12.6	1115	9	CL051289 CH216-73K
C 4	37.4	12.5	721	9	CE726528 tigr-gss-
C 5	37.2	12.4	1043	8	CC314654 FAM32-17J
C 6	37	12.3	675	1	AI857124 MBNAad-23
C 7	37	12.3	1114	9	CL092596 ISB1-22D2
C 8	36.4	12.1	731	9	CG226877 OG0C203TH
C 9	36.4	12.1	819	9	CG360051 OG0E260TH
C 10	36.4	12.1	849	9	CG222409 OG0F036TH
C 11	36.4	12.1	884	8	CC408141 PUHIF57TD
C 12	36.4	12.1	935	8	CC408138 PUHIF57TB
C 13	36.4	12.1	985	9	CL991997 ZMWBHf000
C 14	36.4	12.1	1027	9	CL987077 ZMWBHf000
C 15	36.4	12.1	1058	9	CL985353 ZMWBHf000
C 16	36.2	12.1	728	9	CL368806 RPI144_28
C 17	36	12.0	723	8	AQ795241 nbxb0055N
C 18	35.8	11.9	478	8	B72588 RPI111-9J17
C 19	35.8	11.9	631	9	CE290785 tigr-gss-
C 20	35.8	11.9	682	9	AG175765 Pan trogl
C 21	35.6	11.9	293	8	AZ757570 ew09h01.f
C 22	35.6	11.9	468	2	BF604634 270544 MA
C 23	35.6	11.9	686	2	BB620693 BB620693
C 24	35.4	11.8	669	9	CR041049 Reverse s

ALIGNMENTS

RESULT 1  
CR231209/c

LOCUS  
DEFINITION

ACCESSION  
CR231209

VERSION  
CR231209.1

KEYWORDS  
GSS; Genome survey sequence; MICER.

SOURCE  
Mus musculus

ORGANISM  
Mus musculus

REFERENCE  
1 (bases 1 to 737)

AUTHORS  
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.

TITLE  
Direct Submission

JOURNAL  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES  
source

1..737  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN114f13"  
/clone\_lib="MHPN"

ORIGIN

Query Match 14.5%; Score 43.4; DB 9; Length 737;  
Best Local Similarity 60.0%; Pred. No. 0.11;  
Matches 93; Conservative 0; Mismatches 56; Indels 6; Gaps 1;

QY 1 AAGGAGGAGGAGTGGCTAGGAGATCGAGAGATCAGAAATTTTAACTCAGCCGCGCA 60  
Db 152 AAGGATGCTACAGAGAGAGGAGAAATCAAGATACCTGAGTGGAGAGACTCCGCTCAGCCA 93

QY 61 TTAACATGCTCAAGTACTCTTCATCATATTTCTTAAGAGACACAGTTTCACTCAATGAAT 120  
Db 92 CTGGCAGCATGTGAGAGCGCCACTCTTTTAT-----CAAGAGTCTGCTGAATGAAT 39

QY 121 TCTAAGGCTTTGGGTTTTTATTCAGTGTGCTTCG 155  
Db 38 TCTAAGGCTTTTCAGATTTTATCGCTTGATCCGG 4

RESULT 2  
AZ446937/c

25 35.2 11.7 896 9 CG947926 CG947926  
26 35.2 11.7 899 8 CG965704 CG965704  
27 35.2 11.7 1024 8 CC220759 CC220759  
C 28 35 11.7 395 2 BE092056 BE092056  
C 29 35 11.7 527 2 BE092021 BE092021  
C 30 35 11.7 593 7 CN887426 CN887426  
C 31 35 11.7 597 2 BE091833 BE091833  
C 32 35 11.7 619 7 CN887915 CN887915  
C 33 35 11.7 931 5 BQ890869 BQ890869  
C 34 35 11.7 8596 8 AQ839804 AQ839804  
35 34.8 11.6 649 9 AG277081 AG277081  
36 34.8 11.6 771 8 BZ154035 BZ154035  
37 34.8 11.6 861 8 BZ101919 BZ101919  
C 38 34.6 11.5 476 8 AZ124334 AZ124334  
C 39 34.6 11.5 515 8 AZ124333 AZ124333  
C 40 34.6 11.5 763 8 BZ046496 BZ046496  
C 41 34.6 11.5 1024 9 CNS005BT CNS005BT  
42 34.6 11.5 1113 3 CR705775 CR705775  
C 43 34.4 11.5 543 1 AU261519 AU261519  
44 34.4 11.5 635 5 BU001103 BU001103  
45 34.4 11.5 741 9 AG571746 AG571746

CR231209 737 bp DNA linear GSS 06-JUL-2004  
Reverse strand read from insert in 5'HPRT insertion targeting and  
chromosome engineering clone MHPN114f13, genomic survey sequence.

CR231209.1 GI:50010058  
GSS; Genome survey sequence; MICER.

Mus musculus  
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICER

Location/Qualifiers  
1..737

/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN114f13"  
/clone\_lib="MHPN"

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AZ446937 648 bp DNA linear GSS 04-OCT-2000  
1M0243N04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0243N04 R, genomic survey sequence.  
AZ446937  
AZ446937.1 GI:10598239  
GSS.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 648)  
Dunn,D., Aoyagi,A., Barber,M., Bescorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,  
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von  
Niederhausern,A. and Wright,D. Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0243 row: N column: 04  
Seq primer: CACACGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 648.

FEATURES  
source  
1. .648  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0243N04"  
/sex="Male"  
/lab\_hosts="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptored DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWD42 [gi|4732114|gb|AF129072.1], a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptored mouse DNA was annealed to  
adaptored vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

Query Match 13.0%; Score 39; DB 8; Length 648;  
Best Local Similarity 54.5%; Pred. No. 2.1;  
Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 148 TGCCTCTGCTAGTTTCTGAGGAATCTAAGGCACAACTGAGGAATGAACTCAGGCTTTCCA 207  
DB 313 TGGTCTCTGATGTTTGTGCGAGGAATAGAAACCTGACTGAGACAAATCTGAAGGTTTCA 254  
QY 208 ATTCCGGAATACTCTCCACTGCTTACTGCTCCCTTGGAAATTAGAAGGAGGCAG 267  
DB 253 ATTCAAGTGTGCTGCCAGAACGCTAAATTCATATATGTCAGCATTTCTGAAGACCCAC 194

QY 268 GAGCATAGCTGCCATACACAGGG 290  
DB 193 AAGTATGCTGTTTATAACAAGG 171

RESULT 3  
CL051289 1115 bp DNA linear GSS 31-DEC-2003  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CH216-73K13\_Sp6.1 CH216 Xenopus tropicalis genomic clone  
CH216-73K13, genomic survey sequence.  
CL051289  
CL051289.1 GI:40507202  
GSS.  
Xenopus tropicalis (western clawed frog)  
Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
1 (bases 1 to 1115)  
Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,  
Mardis,E. and Wilson,R.  
A physical map of the xenopus tropicalis genome  
Unpublished (2003)  
Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@watson.wustl.edu  
Insert Length: 175000 Std Error: 0.00  
Seq primer: Sp6 ATTTAGGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 67  
High quality sequence stop: 334.  
Location/Qualifiers  
1. .1115  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-73K13"  
/sex="male"  
/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
BAC library"

Query Match 12.6%; Score 37.8; DB 9; Length 1115;  
Best Local Similarity 54.7%; Pred. No. 5.4;  
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 6 CGAGGCAGTGGGCTAGGGAATCGAGAGATCAGAAATTTTAACTCAGCCAGCCATTAAAC 65  
DB 265 GAGGTCTCTGGCCAAATGTAACCTTTTATGGAACCTTTTACGCCGCCCAACCATCCCA 324  
QY 66 ATGCCCTCAAGTACTCTCATATTTTGAAGACACAGTTCACTGAATGAATCTAA 125  
DB 325 TGGGTTTACATAATTAGACCAATTTTGTGAAAAATCTACACATCCCTGAAATAAACTTAA 384  
QY 126 GGTCTTTGGGTTTAT 142  
DB 385 AGCCTTTCATTTTTTT 401

RESULT 4  
CE726528/c 721 bp DNA linear GSS 30-SEP-2003  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE

CE726528  
tigr-gss-dog-17000329818790 Dog Library Canis familiaris genomic,  
genomic survey sequence.  
CE726528  
CE726528.1 GI:37066648  
GSS.  
Canis familiaris (dog)



Qy 271 CATAGTGCATACCAGGGATGAATTC 299  
|||||  
db 410 CATAGTTGGGATGAAAAGAGACAAACGTC 382

## REFERENCE AUTHORS

55 GTAGTTTCTGAGGAATCTAAGGCACAACGTGAGGAATCAAGTCAGGCTTCCAA 208  
|| |||||  
67 GTGGTTTTCATATAATTGTATTGAAAAAAAACAATGCTGAAGCAAGAGTTTAAA 420

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 819)

1 (bases 1 to 819)  
Whitelaw, C.A., Quackenbush, J., van Aken, S., Utterback, T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other\_GSSs: OG0E260TV  
Contact: Cathy Whitelaw  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

source  
1. .819  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBMa0703124"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 12.1%; Score 36.4; DB 9; Length 819;  
Best Local Similarity 50.6%; Pred. No. 13;  
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 35 TCAGATTTTAAACTCAGCCAGCCATTAACATGCTCAAGTACTCCTATCATATTGTA 94  
Db 262 TCTTTATAATTCACATGATCTATTTCCCTATGCGACATTTTTTTTAAATTGTATA 321  
QY 95 AGAGACAACAGTTCACTGAATGAATCTAAGGCTCTTTGGGTTTTTATCAGTGTCTTCT 154  
Db 322 AGACTGAATTTTCTGTTAATGTAACATAATTTGTTGTTCTTTTGTATTGTCTATA 381  
QY 155 GTAGTTTCTGAGAAATCTAAGGCACAACCTGAGGAATGAAGTCAGGCTTTCCAA 208  
Db 382 GTGGTTTCATAATAATTGTTATGAAAAAACAACTGCTGAAGCAAGAGTTTTAA 435

## RESULT 10

CG222409  
LOCUS  
DEFINITION  
OG0FU36TH ZM 0.7 1.5 KB Zea mays genomic clone ZMBMa0708F23,  
genomic survey sequence.  
CG222409  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

849 bp DNA linear GSS 22-AUG-2003  
CG222409  
Accession  
CG222409.1 GI:34122297  
GSS.  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

1 (bases 1 to 849)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Other\_GSSs: OG0FU36TV  
Contact: Cathy Whitelaw  
TIGR

## TITLE

## JOURNAL

## COMMENT

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

source  
1. .849  
Location/Qualifiers  
/organism="Zea mays"

/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBMa0708F23"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 12.1%; Score 36.4; DB 9; Length 849;  
Best Local Similarity 50.6%; Pred. No. 13;  
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 35 TCAGATTTTAAACTCAGCCAGCCATTAACATGCTCAAGTACTCCTATCATATTGTA 94  
Db 197 TCTTTATAATTCACATGATCTATTTCCCTATGCGACATTTTTTTTAAATTGTATA 256  
QY 95 AGAGACAACAGTTCACTGAATGAATCTAAGGCTCTTTGGGTTTTTATCAGTGTCTTCT 154  
Db 257 AGACTGAATTTTCTGTTAATGTAACATAATTTGTTGTTCTTTTGTATTGTCTATA 316  
QY 155 GTAGTTTCTGAGAAATCTAAGGCACAACCTGAGGAATGAAGTCAGGCTTTCCAA 208  
Db 317 GTGGTTTCATAATAATTGTTATGAAAAAACAACTGCTGAAGCAAGAGTTTTAA 370

## RESULT 11

CC408141  
LOCUS  
DEFINITION  
PUHIF57TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBMa457I18,  
genomic survey sequence.  
CC408141  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

884 bp DNA linear GSS 19-MAY-2003  
CC408141  
Accession  
CC408141.1 GI:30888231  
GSS.  
Zea mays  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

1 (bases 1 to 884)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other\_GSSs: PUHIF57TB  
Contact: Cathy Whitelaw  
TIGR

## TITLE

## JOURNAL

## COMMENT

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

source  
1. .884  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMBMa457I18"  
/clone\_lib="ZM 0.6 1.0 KB"  
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
CoT selected genomic DNA library"

## ORIGIN

Query Match 12.1%; Score 36.4; DB 8; Length 884;  
Best Local Similarity 50.6%; Pred. No. 13;  
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;  
QY 35 TCAGATTTTAAACTCAGCCAGCCATTAACATGCTCAAGTACTCCTATCATATTGTA 94  
Db 389 TCTTTATAATTCACATGATCTATTTCCCTATGCGACATTTTTTTTAAATTGTATA 448

```

QY 95 AGAGACACAGTTCACTGAAATGAAATCTTAAGGTCCTTTGGGTTTTTATCAGTGTCTTCT 154
Db 449 AGACTGAAATTTCTGTTTAATGTTACATTAATTTGTTGTTCTTTTGTATTTGCTATA 508
QY 155 GTAGTTTCTGAGGAATCTTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208
Db 509 GTGTTTTCATAATAATTGTTATTGAAAAAACAATGCTGAAGCAAGAGTTTTAAA 562

RESULT 12
CC408138/c
LOCUS
DEFINITION
  CC408138 935 bp DNA linear GSS 19-MAY-2003
  PUHIF57TB ZM 0.6-1.0 kb Zea mays genomic clone ZMMBta457118,
  genomic survey sequence.
ACCESSION
  CC408138
VERSION
  CC408138.1 GI:30888228
KEYWORDS
  GSS.
SOURCE
  Zea mays
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 935)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other GSSs: PUHIF57TD
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@cigr.org
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..935
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBta457118"
    /clone_lib="ZM 0.6-1.0 KB"
    /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"

FEATURES
  source
    1..935
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /cultivar="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBta457118"
    /tissue_type="immature ear"
    /dev_stage="6-8 weeks"
    /lab_host="DH10B"
    /clone_lib="ZMMBHF"
    /note="Vector: TOPOpcr4; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN
  Query Match 12.1%; Score 36.4; DB 8; Length 935;
  Best Local Similarity 50.6%; Pred. No. 13;
  Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

  QY 35 TCAGAAATTTAACTCAGCCAGCATTAACATGCTCAAGTACTCCTATCATATTGTA 94
  Db 907 TCTTTATAATTCACCTCAATGATCTATTTCCCTATGGCACATTTTTTTAAATGTA 848
  QY 95 AGAGACACAGTTCACTGAAATGAAATCTTAAGGTCCTTTGGGTTTTTATCAGTGTCTTCT 154
  Db 847 AGACTGAAATTTCTGTTTAATGTTAAACATAATTTGTTGTTCTTTTGTATTTGCTATA 788
  QY 155 GTAGTTTCTGAGGAATCTTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208
  Db 787 GTGGTTTCATAATAATTGTTATTGAAAAAACAATGCTGAAGCAAGAGTTTTAAA 734

RESULT 13
CL991997/c
LOCUS
DEFINITION
  CL991997 985 bp DNA linear GSS 23-SEP-2004
  ZMMBHF0003d22.f ZMMBHF Zea mays genomic clone ZMMBHF0003d22 5',
  genomic survey sequence.
ACCESSION
  CL991997
  ZMMBHF0003d22.f
  ZMMBHF Zea mays genomic clone ZMMBHF0003d22 5',
  genomic survey sequence.
  CL991997

QY 95 AGAGACACAGTTCACTGAAATGAAATCTTAAGGTCCTTTGGGTTTTTATCAGTGTCTTCT 154
Db 449 AGACTGAAATTTCTGTTTAATGTTACATTAATTTGTTGTTCTTTTGTATTTGCTATA 508
QY 155 GTAGTTTCTGAGGAATCTTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208
Db 509 GTGTTTTCATAATAATTGTTATTGAAAAAACAATGCTGAAGCAAGAGTTTTAAA 562

RESULT 14
CL987077/c
LOCUS
DEFINITION
  CL987077 1027 bp DNA linear GSS 23-SEP-2004
  ZMMBHe0003k03.f ZMMBHe Zea mays genomic clone ZMMBHe0003k03 5',
  genomic survey sequence.
ACCESSION
  CL987077
VERSION
  CL987077.1 GI:52555155
KEYWORDS
  GSS.
SOURCE
  Zea mays
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 1027)
  Ma,J., SanMiguel,P., Liu,R., Haller,K., Soderlund,C. and
  Bennetzen,J.
  ZMMBH sequences
  Unpublished (2004)
  Contact: Jeff Bennetzen
  Bennetzen Lab

QY 35 TCAGAAATTTAACTCAGCCAGCATTAACATGCTCAAGTACTCCTATCATATTGTA 94
Db 567 TCTTTATAATTCACCTCAATGATCTATTTCCCTATGGCACATTTTTTTAAATGTA 508
QY 95 AGAGACACAGTTCACTGAAATGAAATCTTAAGGTCCTTTGGGTTTTTATCAGTGTCTTCT 154
Db 507 AGACTGAAATTTCTGTTTAATGTTAAACATAATTTGTTGTTCTTTTGTATTTGCTATA 448
QY 155 GTAGTTTCTGAGGAATCTTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCAA 208
Db 447 GTGGTTTCATAATAATTGTTATTGAAAAAACAATGCTGAAGCAAGAGTTTTAAA 394

RESULT 15
CL991997/c
LOCUS
DEFINITION
  CL991997 985 bp DNA linear GSS 23-SEP-2004
  ZMMBHF0003d22.f ZMMBHF Zea mays genomic clone ZMMBHF0003d22 5',
  genomic survey sequence.
ACCESSION
  CL991997
  ZMMBHF0003d22.f
  ZMMBHF Zea mays genomic clone ZMMBHF0003d22 5',
  genomic survey sequence.
  CL991997

```



***This Page Blank (uspto)***



```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152883
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-152883

Query Match      11.4%; Score 34.2; DB 4; Length 601;
Best Local Similarity 60.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 57; Conservative

QY 103 CAGTTCACCTGAAATGAATCTTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTC 162
Db 550 CAATTCCTCTAAAGTTCACCTCTACCGTTGTAATTTTACACTAAGTGGGCTTAGATAGTTTA 491

QY 163 TGAGGAATCTAAGGCACAACTGAGGAATGAAGTC 197
Db 490 CCAAGAATCTTAAGGAACATAGCTTCAATGAATC 456

RESULT 3
US-09-949-016-152884/c
; Sequence 152884, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152884
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-152884

Query Match      11.4%; Score 34.2; DB 4; Length 601;
Best Local Similarity 60.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 57; Conservative

QY 103 CAGTTCACCTGAAATGAATCTTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTC 162
Db 226 CAATTCCTCTAAAGTTCACCTCTACCGTTGTAATTTTACACTAAGTGGGCTTAGATAGTTTA 167

QY 163 TGAGGAATCTAAGGCACAACTGAGGAATGAAGTC 197
Db 166 CCAAGAATCTTAAGGAACATAGCTTCAATGAATC 132

RESULT 4
US-09-949-016-152885/c
; Sequence 152885, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152885
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-152885

Query Match      11.4%; Score 34.2; DB 4; Length 601;
Best Local Similarity 60.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 57; Conservative

QY 103 CAGTTCACCTGAAATGAATCTTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTC 162
Db 195 CAATTCCTCTAAAGTTCACCTCTACCGTTGTAATTTTACACTAAGTGGGCTTAGATAGTTTA 136

QY 163 TGAGGAATCTAAGGCACAACTGAGGAATGAAGTC 197
Db 135 CCAAGAATCTTAAGGAACATAGCTTCAATGAATC 101

RESULT 5
US-09-949-016-16038/c
; Sequence 16038, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16038
; LENGTH: 166698
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(166698)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16038

Query Match      11.4%; Score 34.2; DB 4; Length 166698;
Best Local Similarity 60.0%; Pred. No. 3.1; Mismatches 0; Indels 0; Gaps 0;
Matches 57; Conservative

QY 103 CAGTTCACCTGAAATGAATCTTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTTC 162
Db 3104 CAATTCCTCTAAAGTTCACCTCTACCGTTGTAATTTTACACTAAGTGGGCTTAGATAGTTTA 3045

QY 163 TGAGGAATCTAAGGCACAACTGAGGAATGAAGTC 197
Db 3044 CCAAGAATCTTAAGGAACATAGCTTCAATGAATC 3010

```

```
RESULT 6
US-09-949-016-12683
; Sequence 12683, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12683
; LENGTH: 152524
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12683

Query Match      11.3%; Score 33.8; DB 4; Length 152524;
Best Local Similarity 50.3%; Pred. No. 4.1;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 12 AGTGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCATTAAACATGCTT 71
Db 2830 AGTGGTGAACAATTTCTTTTGATAACAAGTTCGGTGGCTAGGCAATTAACACT 2889

Qy 72 CAAGTACTCTCATATTTTGTAAAGACAAACAGTTCACCTGAAATGAATTTTAAGGTCTT 131
Db 2890 CCCTCAATCCTTTAAAAATTAAGCAACTTAAGATAGCTTTACAAAAACAAAACACCC 2949

Qy 132 TGGGTTTTTATCAGTGTCTTCTCTAGTCTTCTGAGGAATCTAAG 176
Db 2950 TGTGCTCTATAAGGCAGAACGTGTTCTGTAAAGATGAACCCAAAG 2994

Query Match      11.3%; Score 33.8; DB 4; Length 152524;
Best Local Similarity 50.3%; Pred. No. 4.1;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 12 AGTGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCATTAAACATGCTT 71
Db 2830 AGTGGTGAACAATTTCTTTTGATAACAAGTTCGGTGGCTAGGCAATTAACACT 2889

Qy 72 CAAGTACTCTCATATTTTGTAAAGACAAACAGTTCACCTGAAATGAATTTTAAGGTCTT 131
Db 2890 CCCTCAATCCTTTAAAAATTAAGCAACTTAAGATAGCTTTACAAAAACAAAACACCC 2949

Qy 132 TGGGTTTTTATCAGTGTCTTCTGTAGTCTTCTGAGGAATCTAAG 176
Db 2950 TGTGCTCTATAAGGCAGAACGTGTTCTGTAAAGATGAACCCAAAG 2994

RESULT 7
US-09-949-016-13194
; Sequence 13194, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13194
; LENGTH: 152524
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13194

Query Match      11.3%; Score 33.8; DB 4; Length 152524;
Best Local Similarity 50.3%; Pred. No. 4.1;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 12 AGTGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCATTAAACATGCTT 71
Db 2830 AGTGGTGAACAATTTCTTTTGATAACAAGTTCGGTGGCTAGGCAATTAACACT 2889

Qy 72 CAAGTACTCTCATATTTTGTAAAGACAAACAGTTCACCTGAAATGAATTTTAAGGTCTT 131
Db 2890 CCCTCAATCCTTTAAAAATTAAGCAACTTAAGATAGCTTTACAAAAACAAAACACCC 2949

Qy 132 TGGGTTTTTATCAGTGTCTTCTGTAGTCTTCTGAGGAATCTAAG 176
Db 2950 TGTGCTCTATAAGGCAGAACGTGTTCTGTAAAGATGAACCCAAAG 2994

RESULT 8
US-09-949-016-12637
; Sequence 12637, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12637
; LENGTH: 17763
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12637

Query Match      11.1%; Score 33.4; DB 4; Length 17763;
Best Local Similarity 48.7%; Pred. No. 1.8;
Matches 91; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 71 TCAAGTACTCTCATATTTTGTAAAGACAAACAGTTCACCTGAAATGAATTTTAAGGTCT 130
Db 11584 TTAAAGTCTCTATTTATTATTGTGTGGAGTCTAAGTCTTTTGTAGTCTCTAAGGACT 11643

Qy 131 TTGGGTTTTTATCAGTGTCTTCTGTAGTCTTCTGAGGAATCTAAGGCACAACTGAGGAA 190
Db 11644 TGGTTTTTGAATCTGGGTGCTCTCTGTATTGGGTGCATATATATTAGGATAGTTAACTCT 11703

Qy 191 TGAAGTCAGGCTTTCCAAATCCCGAAATATCTCTCCACTGCTTACTCATGTCCTTGAA 250
Db 11704 TCTTGTGAATGTATGCTCTTTACCATAATAGTGGCTTGTCTCTTTGTATCTTTGTTG 11763

Qy 251 ATTAAGA 257
Db 11764 GTTTAA 11770

RESULT 9
US-09-949-016-13143
; Sequence 13143, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
Qy 12 AGTGGCTAGGAGATCGAGAGATCAGAAATTTTAAACTCAGCCAGCCATTAAACATGCTT 71
Db 2830 AGTGGTGAACAATTTCTTTTGATAACAAGTTCGGTGGCTAGGCAATTAACACT 2889

Qy 72 CAAGTACTCTCATATTTTGTAAAGACAAACAGTTCACCTGAAATGAATTTTAAGGTCTT 131
Db 2890 CCCTCAATCCTTTAAAAATTAAGCAACTTAAGATAGCTTTACAAAAACAAAACACCC 2949

Qy 132 TGGGTTTTTATCAGTGTCTTCTCTAGTCTTCTGAGGAATCTAAG 176
Db 2950 TGTGCTCTATAAGGCAGAACGTGTTCTGTAAAGATGAACCCAAAG 2994
```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13143
; LENGTH: 17764
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13143

Query Match 11.1%; Score 33.4; DB 4; Length 17764;
Best Local Similarity 48.7%; Pred. No. 1.8;
Matches 91; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 71 TCAAGTACTCCTATCATATTTGTAGAGACACAGTTCACCTGAAATGAATTTCTAAGGTCT 130
Db 11584 TTAAGTCTCCTATTAATTAATTTGTGGAGTCTAAGTCTTTTGTAGGTCTCTAAGGACT 11643

Qy 131 TTGGGTTTTATCAGTGTCTTCTGTAGTTCTTGAGAAATCTAAGGCACAACTGAGGAA 190
Db 11644 TGGTTTTGAATCTGGTCTCTCTGTAATTTGGTGCATATATATTTAGGATAGTTAACTCT 11703

Qy 191 TGAAGTCAGGCTTTCATATCCGAATACCTCCACTGCTTACTCATGTCCTTGGAA 250
Db 11704 TCTTGTGAATTGATCCCTTTACCATAATGATAGTGGCCTTGCTCTTTTGATCTTTGTG 11763

Qy 251 ATTAAGA 257
Db 11764 GTTTAA 11770

RESULT 10
US-08-110-158-5/c
; Sequence 5, Application US/08110158
; Patent No. 5605821
; GENERAL INFORMATION:
; APPLICANT: McEver, Rodger P.
; APPLICANT: Pan, Junliang
; TITLE OF INVENTION: Expression Control Sequences of the
; TITLE OF INVENTION: P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,158
; FILING DATE: 19930820
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/320,408
; FILING DATE: 08-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4866 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
US-08-110-158-5

Query Match 10.9%; Score 32.6; DB 1; Length 4866;
Best Local Similarity 57.3%; Pred. No. 1.7;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 18 CTAGGAGATCGAGATCAGATTTTAACTCAGCCAGCCCAATTAACATGCCTCAAGTA 77
Db 3105 CTATCAGCATCATGAGAACAGATAATACATCAGCAATGTGATAATAGGATAAAT 3046

Qy 78 CTCTTATCATATTTGTAAAGAGACAAACAGTTCACCTGAAATGAAT 120
Db 3045 CACCAATACATAGACATGTAAGGACCATTAAATGAATTAAT 3003

RESULT 11
US-09-949-016-16142/c
; Sequence 16142, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16142
; LENGTH: 45323
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16142

Query Match 10.9%; Score 32.6; DB 4; Length 45323;
Best Local Similarity 57.3%; Pred. No. 5.5;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 18 CTAGGAGATCGAGATCAGATTTTAACTCAGCCAGCCCAATTAACATGCCTCAAGTA 77
Db 337 CTATCAGCATCATGAGAACAGATAATACATCAGCAATGTGATAATAGGATAAAT 278

Qy 78 CTCTTATCATATTTGTAAAGAGACAAACAGTTCACCTGAAATGAAT 120
Db 277 CACCAATACATAGACATGTAAGGACCATTAAATGAATTAAT 235

RESULT 12
US-09-949-016-15896/c
; Sequence 15896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 15896
; LENGTH: 76472
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(76472)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15896

Query Match          10.9%; Score 32.6; DB 4; Length 76472;
Best Local Similarity 57.3%; Pred. No. 7.3;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 71 TCAAGTACTCCTATCATATTTGTAAGAGACACAGTTCTCACTGAAATGAATTTCTAAGGTCT 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30938 TTAAGTCTCCCATTTATTGTTGGAGTCTTAATTTCTTTGTAGGTCTCTAAGGACT 30879

QY 131 TTGGGTTTTTATCAGTGTCTCTCTGTAGTTTCTGAGGAAATCT 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30878 TGCATTATGAATCCGGGTCTCTCTGTATTGGGTGAATATATAT 30836

RESULT 13
US-09-949-016-17484
; Sequence 17484, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17484
; LENGTH: 92155
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17484

Query Match          10.9%; Score 32.6; DB 4; Length 92155;
Best Local Similarity 57.3%; Pred. No. 8.1;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 71 TCAAGTACTCCTATCATATTTGTAAGAGACACAGTTCTCACTGAAATGAATTTCTAAGGTCT 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72852 TTAAGTCTCCCATTTATTGTTGGAGTCTTAAGTCTCTTTATAGGTCTCTAAGGACT 72911

QY 131 TTGGGTTTTTATCAGTGTCTCTCTGTAGTTTCTGAGGAAATCT 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72912 TGCATTATGAATCGGTGCTCTCTGTATTGGGTGCATATATAT 72954

RESULT 14
US-09-949-016-12166/c
; Sequence 12166, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12166
; LENGTH: 92276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(92276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12166

Query Match          10.9%; Score 32.6; DB 4; Length 92276;
Best Local Similarity 57.3%; Pred. No. 8.1;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 71 TCAAGTACTCCTATCATATTTGTAAGAGACACAGTTCTCACTGAAATGAATTTCTAAGGTCT 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46742 TTAAGTCTCCCATTTATTGTTGGAGTCTTAATTTCTTTGTAGGTCTCTAAGGACT 46683

QY 131 TTGGGTTTTTATCAGTGTCTCTCTGTAGTTTCTGAGGAAATCT 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46682 TGCATTATGAATCCGGGTCTCTCTGTATTGGGTGAATATATAT 46640

RESULT 15
US-09-949-016-17009
; Sequence 17009, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17009
; LENGTH: 205163
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17009

Query Match          10.9%; Score 32.6; DB 4; Length 205163;
Best Local Similarity 55.9%; Pred. No. 12;
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 71 TCAAGTACTCCTATCATATTTGTAAGAGACACAGTTCTCACTGAAATGAATTTCTAAGGTCT 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158979 TTAAGTCTCCCATTTATTGTTGGAGTCTTAAGTCTCTTTGTAGATCTCTAAGGACT 159038

QY 131 TTGGGTTTTTATCAGTGTCTCTCTGTAGTTTCTGAGGAAATCTAAGGACA 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159039 TGCATTATGAATCTGGGTCTCTCTGTATTGGGTGCATATATCTTTAGGACA 159089

Search completed: August 4, 2005, 14:40:55
Job time : 67.8219 secs
```

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 535.142 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-5

Perfect score: 300

Sequence: 1 aaggaggagcgctggcta.....ataaccggatgaacttct 300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10J\_PUBCOMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	US-09-899-276-5	Sequence 5, Appli
2	299.6	99.9	934	US-10-027-632-121412	Sequence 121412,
3	299.6	99.9	934	US-10-027-632-121412	Sequence 121412,
4	296.8	98.9	3221	US-10-685-705-3	Sequence 3, Appli
5	296.8	98.9	11793	US-10-685-705-4	Sequence 4, Appli
6	204	68.0	5926	US-10-311-455-1625	Sequence 1625, Ap
7	194.4	64.8	5926	US-10-311-455-1626	Sequence 1626, Ap

C 8	40.4	13.5	563	13	US-10-027-632-129675	Sequence 129675,
C 9	40.4	13.5	563	17	US-10-027-632-129675	Sequence 129675,
C 10	36.8	12.3	204521	13	US-10-087-192-958	Sequence 958, App
C 11	36.8	12.3	290367	20	US-10-719-993-6887	Sequence 6887, Ap
C 12	35.2	11.7	403	13	US-10-027-632-89184	Sequence 89184, A
C 13	35.2	11.7	403	17	US-10-027-632-89184	Sequence 89184, A
C 14	35.2	11.7	2286	13	US-10-027-632-102218	Sequence 102218,
C 15	35.2	11.7	2286	17	US-10-027-632-102218	Sequence 102218,
C 16	35	11.5	38855	11	US-09-973-278-930	Sequence 930, App
C 17	34.4	11.5	1691139	14	US-10-067-514-1	Sequence 1, Appli
C 18	34.4	11.5	1691139	17	US-10-419-723-1	Sequence 1, Appli
C 19	34.4	11.5	1691140	22	US-10-868-397-1	Sequence 1, Appli
C 20	34.2	11.4	549	22	US-10-972-079-71139	Sequence 71139, A
C 21	34.2	11.4	618	13	US-10-027-632-199837	Sequence 199837,
C 22	34.2	11.4	618	13	US-10-027-632-199838	Sequence 199838,
C 23	34.2	11.4	618	17	US-10-027-632-199837	Sequence 199837,
C 24	34.2	11.4	618	17	US-10-027-632-199838	Sequence 199838,
C 25	34.2	11.4	5923	21	US-10-956-157-4561	Sequence 4561, Ap
C 26	34	11.3	1924	13	US-10-027-632-104612	Sequence 104612,
C 27	34	11.3	1924	17	US-10-027-632-104612	Sequence 104612,
C 28	33.6	11.2	693	17	US-10-282-122A-18166	Sequence 18166, A
C 29	33.2	11.1	1483	22	US-10-764-212-79	Sequence 79, Appli
C 30	33.2	11.1	344548	13	US-10-087-192-334	Sequence 334, App
C 31	33	11.0	600	22	US-10-972-079-9597	Sequence 9597, Ap
C 32	33	11.0	600	22	US-10-972-079-9598	Sequence 9598, Ap
C 33	32.8	10.9	524	20	US-10-425-115-168400	Sequence 168400,
C 34	32.8	10.9	2893	13	US-10-027-632-111626	Sequence 111626,
C 35	32.8	10.9	2893	17	US-10-027-632-111626	Sequence 111626,
C 36	32.8	10.9	17918	17	US-10-221-613-382	Sequence 382, App
C 37	32.6	10.9	201	19	US-10-741-601-20010	Sequence 20010, A
C 38	32.6	10.9	201	21	US-10-741-600-55435	Sequence 55435, A
C 39	32.6	10.9	3196	10	US-09-764-891-6999	Sequence 6999, Ap
C 40	32.6	10.9	53323	19	US-10-741-601-5695	Sequence 5695, Ap
C 41	32.6	10.9	53323	21	US-10-741-600-17769	Sequence 17769, A
C 42	32.6	10.9	78025	14	US-10-020-141-9	Sequence 9, Appli
C 43	32.6	10.9	105219	13	US-10-087-192-658	Sequence 658, App
C 44	32.4	10.8	600	22	US-10-972-079-9599	Sequence 9599, Ap
C 45	32.4	10.8	84073	13	US-10-087-192-712	Sequence 712, App

#### ALIGNMENTS

RESULT 1  
US-09-899-276-5  
; Sequence 5, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Dellus, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899,276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-5

Query Match 100.0%; Score 300; DB 9; Length 300;  
Best Local Similarity 100.0%; Pred. No. 5e-88;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAAACTGAGCCAGCCA 60  
 Db 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAAACTGAGCCAGCCA 60  
 QY 61 TTAACATGCTCAAGTACTCCTATCATATTTGTAAGAGACAACTGTTCACTGAATGAAT 120  
 Db 61 TTAACATGCTCAAGTACTCCTATCATATTTGTAAGAGACAACTGTTCACTGAATGAAT 120  
 QY 121 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 180  
 Db 121 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 180  
 QY 181 AACTGAGGAATGAAGTCAAGCTTCCAAATTCGCCGAATCTCCTCCACTGCTTACTCATG 240  
 Db 181 AACTGAGGAATGAAGTCAAGCTTCCAAATTCGCCGAATCTCCTCCACTGCTTACTCATG 240  
 QY 241 TCCCTTGGAAATTAAGAGGAAGCCAGGAGATAGCTGCCATAACCCAGGATGAATCTCT 300  
 Db 241 TCCCTTGGAAATTAAGAGGAAGCCAGGAGATAGCTGCCATAACCCAGGATGAATCTCT 300  
 RESULT 2  
 US-10-027-632-121412  
 ; Sequence 121412, Application US/10027632  
 ; Publication No. US20020198371A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: US 60/185,218  
 ; PRIOR FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/167,363  
 ; PRIOR FILING DATE: 1999-11-23  
 ; PRIOR APPLICATION NUMBER: US 60/156,358  
 ; PRIOR FILING DATE: 1999-09-28  
 ; PRIOR APPLICATION NUMBER: US 60/146,002  
 ; PRIOR FILING DATE: 1999-08-09  
 ; NUMBER OF SEQ ID NOS: 325720  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 121412  
 ; LENGTH: 934  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-10-027-632-121412  
 Query Match 99.9%; Score 299.6; DB 13; Length 934;  
 Best Local Similarity 99.7%; Pred. No. 1.2e-87;  
 Matches 299; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAAACTGAGCCAGCCA 60  
 Db 85 AAGGAGGAGGAGTGGGCTAGGAGATCGAGAGATCAGAAATTTAAACTGAGCCAGCCA 144  
 QY 61 TTAACATGCTCAAGTACTCCTATCATATTTGTAAGAGACAACTGTTCACTGAATGAAT 120  
 Db 145 TTAACATGCTCAAGTACTCCTATCATATTTGTAAGAGACAACTGTTCACTGAATGAAT 204  
 QY 121 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 180  
 Db 205 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 264  
 QY 181 AACTGAGGAATGAAGTCAAGCTTCCAAATTCGCCGAATCTCCTCCACTGCTTACTCATG 240  
 Db 240 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 300  
 QY 181 AACTGAGGAATGAAGTCAAGCTTCCAAATTCGCCGAATCTCCTCCACTGCTTACTCATG 240  
 Db 240 TCTAAGGCTCTTTGGGTTTTTATCAGTGTGCTTCTGTAGTTCTGAGGAAATCTAAGGCAC 300  
 RESULT 4  
 US-10-685-705-3  
 ; Sequence 3, Application US/10685705  
 ; Publication No. US20040177387A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: University of Kentucky Research Foundation  
 ; APPLICANT: JAYAKRISHNA, Ambati  
 ; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macu

;/ TITLE OF INVENTION: Degeneration  
;/ FILE REFERENCE: 050229-0415  
;/ CURRENT APPLICATION NUMBER: US/10/685,705  
;/ CURRENT FILING DATE: 2003-10-16  
;/ PRIOR APPLICATION NUMBER: 60/422,096  
;/ PRIOR FILING DATE: 2002-10-30  
;/ NUMBER OF SEQ ID NOS: 10  
;/ SOFTWARE: PatentIn version 3.2  
;/ SEQ ID NO 3  
;/ LENGTH: 3221  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
US-10-685-705-3

Query Match 98.9%; Score 296.8; DB 19; Length 3221;  
Best Local Similarity 99.3%; Pred. No. 1.7e-86;  
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 AAGGAGGAGGCGAGTGGGCTAGGAGATCGAGATCAGAAATTTTAAACTCAGCCCGCA 60  
DB 1445 AAGGAGGAGGCGAGTGGGCTAGGAGATCGAGATCAGAAATTTTAAACTCAGCCCGCA 1504  
  
QY 61 TTAACATGCCTCAAGTACTCCTCATATTTGTAAGACACACAGTTCACCTGAATGAAT 120  
DB 1505 TTAACATGCCTCAAGTACTCCTCATATTTGTAAGACACACAGTTCACCTGAATGAAT 1564  
  
QY 121 TCTAAGGCTTTGGGTTTTTATCAGTGTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180  
DB 1565 TCTAAGGCTTTGGGTTTTTATCAGTGTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 1624  
  
QY 181 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCCGAAATACCTCTCCACTGCTTACTCATG 240  
DB 1625 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCCGAAATACCTCTCCACTGCTTACTCATG 1684  
  
QY 241 TCCCTTGGAAATTAAGAAGGAGCCAGGAGATAGCTGCCATAACCCAGGGATGAATCTCT 300  
DB 1685 TCCCTTGGAAATTAAGAAGGAGCCAGGAGATAGCTGCCATAACCCAGGGATGAATCTCT 1744

RESULT 5  
US-10-685-705-4  
;/ Sequence 4, Application US/10685705  
;/ Publication No. US20040177387A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: University of Kentucky Research Foundation  
;/ APPLICANT: JAYAKRISHNA, Ambati  
;/ TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
;/ TITLE OF INVENTION: Degeneration  
;/ FILE REFERENCE: 050229-0415  
;/ CURRENT APPLICATION NUMBER: US/10/685,705  
;/ CURRENT FILING DATE: 2003-10-16  
;/ PRIOR APPLICATION NUMBER: 60/422,096  
;/ PRIOR FILING DATE: 2002-10-30  
;/ NUMBER OF SEQ ID NOS: 10  
;/ SOFTWARE: PatentIn version 3.2  
;/ SEQ ID NO 4  
;/ LENGTH: 11793  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
US-10-685-705-4

Query Match 98.9%; Score 296.8; DB 19; Length 11793;  
Best Local Similarity 99.3%; Pred. No. 3.2e-86;  
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGGAGGAGGCGAGTGGGCTAGGAGATCGAGATCAGAAATTTTAAACTCAGCCCGCA 60  
DB 6249 AAGGAGGAGGCGAGTGGGCTAGGAGATCGAGATCAGAAATTTTAAACTCAGCCCGCA 6308  
  
QY 61 TTAACATGCCTCAAGTACTCCTCATATTTGTAAGACACACAGTTCACCTGAATGAAT 120  
DB 6309 TTAACATGCCTCAAGTACTCCTCATATTTGTAAGACACACAGTTCACCTGAATGAAT 6368

QY 121 TCTAAGGCTTTGGGTTTTTATCAGTGTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180  
DB 6369 TCTAAGGCTTTGGGTTTTTATCAGTGTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 6428  
  
QY 181 AACTGAGGAATGAAGTCAGGCTTTCCAAATTCGGAATACCTCTCCACTGCTTACTCATG 240  
DB 6429 AACTGAGGAATGAAGTCAGGCTTTCCAAATTCGGAATACCTCTCCACTGCTTACTCATG 6488  
  
QY 241 TCCCTTGGAAATTAAGAAGGAGCCAGGAGATAGCTGCCATAACCCAGGGATGAATCTCT 300  
DB 6489 TCCCTTGGAAATTAAGAAGGAGCCAGGAGATAGCTGCCATAACCCAGGGATGAATCTCT 6548

RESULT 6  
US-10-311-455-1625  
;/ Sequence 1625, Application US/10311455  
;/ Publication No. US20030143606A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: OLEK, Alexander  
;/ APPLICANT: PIEPENBROCK, Christian  
;/ APPLICANT: BERLIN, Kurt  
;/ TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter  
;/ TITLE OF INVENTION: cytosine methylation  
;/ FILE REFERENCE: 5013.1014  
;/ CURRENT APPLICATION NUMBER: US/10/311,455  
;/ CURRENT FILING DATE: 2002-12-16  
;/ PRIOR APPLICATION NUMBER: PCT/EP01/07537  
;/ PRIOR FILING DATE: 2001-07-02  
;/ PRIOR APPLICATION NUMBER: DE 10032529.7  
;/ PRIOR FILING DATE: 2000-06-30  
;/ PRIOR APPLICATION NUMBER: DE 10043826.1  
;/ PRIOR FILING DATE: 2000-09-01  
;/ NUMBER OF SEQ ID NOS: 2424  
;/ SEQ ID NO 1625  
;/ LENGTH: 5926  
;/ TYPE: DNA  
;/ ORGANISM: Artificial Sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1625

Query Match 68.0%; Score 204; DB 15; Length 5926;  
Best Local Similarity 80.0%; Pred. No. 9e-56;  
Matches 240; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1 AAGGAGGAGGCGAGTGGGCTAGGAGATCGAGATCAGAAATTTTAAACTCAGCCCGCA 60  
DB 3693 AAGGAGGAGGCTAGTGGGTTAGGAGATCGAGATTTTAAATTTAGTTAGTTA 3752  
  
QY 61 TTAACATGCCTCAAGTACTCCTATCATATTTGTAAAGACACACAGTTCACCTGAAATGAAT 120  
DB 3753 TTAATATGTTTAAAGTATTTTATATATTTTGTAAAGATAATAGTATTATGAAATGAAT 3812  
  
QY 121 TCTAAGGCTTTGGGTTTTTATCAGTGTCTTCTGTAGTTTCTGAGGAAATCTAAGGCAC 180  
DB 3813 TTTAAGGCTTTGGGTTTTTATAGTGTGTTTTTGTAGTTTTTGTAGGAAATTTAAGGTAT 3872  
  
QY 181 AACTGAGGAATGAAGTCAGGCTTTCCAAATCCCGAAATACCTCTCCACTGCTTACTCATG 240  
DB 3873 AATTGAGGAATGAAGTTAGGTTTTTAAATTTTTCGAAATATTTTTTATGTTTATTTATG 3932  
  
QY 241 TCCCTTGGAAATTAAGAAGGAGCCAGGAGATAGCTGCCATAACCCAGGGATGAATCTCT 300  
DB 3933 TTTTTCGAAATTAAGAAGGAGTTAGGAGATAGTGTGTTATATATATAGGATGAATTTT 3992

RESULT 7  
US-10-311-455-1626/c  
;/ Sequence 1626, Application US/10311455  
;/ Publication No. US20030143606A1  
;/ GENERAL INFORMATION:  
;/ APPLICANT: OLEK, Alexander  
;/ APPLICANT: PIEPENBROCK, Christian

```

; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determin
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1826
; LENGTH: 5926
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-311-455-1626

```

Query Match	64.88;	Score 194.4;	DB 15;	Length 5926;
Best Local Similarity	78.00;	Pred. No. 1.1e-52;		
Matches 234;	Conservative 0;	Mismatches 66;	Indels 0;	Gaps 0;
QY	1	AAGGAGGAGGCGAGTGGGCTAGGGAATCGAGAGATCAGAAATTTTAAACTCAGCCGACCA	60	
DB	2234	AAAAAAAAACAATAAACTAAAAAAATCGAAAAATCAAAATTTTAAACTCAAGCCCAACCA	2175	
QY	61	TTAATCGCCTCAAGTACTCCTATCATATTTGTAAAGAGACAACAGTTTCACTGAAATGAAT	120	
DB	2174	TTAATCATACCTCAATACTCCTATCATATTTATAAAAACAACAAATTCATTAATAATAAT	2115	
QY	121	TCTAAGGTCCTTTGGGTTTTTATCAGTGTGCTTCGTGTAGTTTCTGAGGAAATCTAAGGCAC	180	
DB	2114	TCTAAATCTTTAAATTTTATCAATATACCTCTATAATTTCTAAAAAAATCTAAAAACAC	2055	
QY	181	AACGTGAGGAATGAAGTCAGGCTTTCCAAATCCCGAAATACCTCCCTCACTGCTTACTCATG	240	
DB	2054	AACATAAAAAATAAATCAAACTTTCCAAATCCCGAAATACCTCTCCCACTCTTACTCATTA	1995	
QY	241	TCCCTTGGAAATTAAGAAGGAAGCCAGGAGCATAGTCGCCATAACCCAGGGATGAACCTTCT	300	
DB	1904	TCCCTTAATAATTAATAAATAAATCAAAATACTTACCATAACCCAAATAATAAATCTCT	1935	

```

RESULT 8
US-10-027-632-129675/c
; Sequence 129675, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 129675
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-129675

```

```
Query Match          13.5%; Score 40.4; DB 13; Length 563;
Best Local Similarity 52.3%; Pred. NO. 0.018;
Matches 137; Conservative 0; Mismatches 121; Indels 4; Gaps 2;
```

Qy	38	GAATTTTAAACTCAGCCCGACCCATTAACTGCCTCAAGTACTCCTATCATATTGTGAAGA	97
Db	471	GAGGTGCCGCAACACTCAATAATCATTTTACCTTCGTGTTTTCATTGTGTAAG	412
Qy	98	GACAACAGTTTCACTGAAATGAAATCTTAAGGTCTTTGGGTTTTTATCAGTGTGCTCTGTGA	157
Db	411	GGAAATACTTGAATAAATAAATGCAACTGCCTCTTTCAACTCAT-TCTATAATATTGTG	353
Qy	158	GTTTCTCAGGAAATCTAAGGCACAACTGAGGAATGAAGTCAGGCTTTCCCAATTCGCCGAAA	217
Db	352	ATCATTAAGGACAGTGTGGAGCAGCTGAAANATGAAGCAAAATCTGCTTCCCAGGA	293
Qy	218	TACTCCTTCCACTGCTTACTCATGTCCCTTGGAAATTAAGAAGGACCGAGAGATAGCT	277
Db	292	TGCTCTCCACCC---TAATTTCACTCTTGGAAACCGAGTAAGAATCTAGAAGGTCTGGT	236
Qy	278	GCCATACCCAGGGATGAAGTTC	299
Db	235	TCCATAGCAGGTATAYATTTC	214

```

RESULT 9
US-10-027-632-129675/c
; Sequence 129675, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108627.129
; CURRENT APPLICATION NUMBER: US/10/027.632

```

```

CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 129675
LENGTH: 563
TYPE: DNA
ORGANISM: Human
US-10-027-632-129675

```

	Query Match	Best Local Similarity	Score	DB 17	Length	563;
	Matches	137; Conservative	52.3%; P: 0.018;	Indels	4;	Gaps
QY	38	GAATTTAACTCAGCCGACCCATTAACTGCTTCAAGTACTCTTATCAATATTTTGTAA	97			
Db	471	GAAGTAGCCGCAACCTCAATAATCAATCTCTCTGTGTGTTTCATTTGTAAG	412			
QY	98	GACAAAGTTCACTGAAATGAATTCCTAAGGCTCTTTGGGTTTTATCACTGTGCTTCTGTA	157			

```
Db 411 GGAATACTTGAATAAATAAATGCAACTGCCTCTTTCAACTCAT-TCTATAATATTGTG 353
QY 158 GTTCTCTGAGGAATCTAAGGCACAACTTGAGGAATGAAGTCAGGCTTTCCAATTTCCCGAAA 217
Db 352 ATCATTAAGACACAGTGGAGCAGAGCTGAAAATGAAGCAGAAATTCGTCTCCCGAGA 293
QY 218 TACTCTCCACTCTTACTCATCTGCCCTTGGAATTAAGAAGGACCGAGCATAGCT 277
Db 292 TGCTCTCCCAACC---TAATTTCACTCTCTTGAGAACCAAGGTAAAGATCTAGAAGTCTGGT 236
QY 278 GCCATAACCAAGGATGAAGTTC 299
Db 235 TCCATAGCAAGGTATATATTC 214

RESULT 10
US-10-087-192-958/c
; Sequence 958, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 958
; LENGTH: 204621
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(204621)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-958

Query Match 12.3%; Score 36.8; DB 13; Length 204621;
Best Local Similarity 58.0%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 42 TTAAACTCAGCCAGCCATTAACTGCTTAAAGTCTTTGGTCTTTATCAGTGTGCTTC 153
Db 105675 TTCACAGTCAACACAGCAATGAAGTTTATGAATTTCTCTTTTATGCTTTTGAAGACATA 105616

QY 102 ACAGTTCACCTGAATGAATCTAAAGTCTTTGGTCTTTATCAGTGTGCTTC 153
Db 105615 ATACTTATGAGACCAATTCAGAAAGTCAATGCGATTTTCAGCAGGGTGAGTC 105564

RESULT 11
US-10-719-993-6887/c
; Sequence 6887, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6887
; LENGTH: 290367
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(290367)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6887

Query Match 12.3%; Score 36.8; DB 20; Length 290367;
Best Local Similarity 58.0%; Pred. No. 5.2;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 42 TTAAACTCAGCCAGCCATTAACTGCTTAAAGTCTTTGGTCTTTATCAGTGTGCTTC 153
Db 202461 TTCACAGTCAACACAGCAATGAAGTTTATGAATTTCTCTTTTATGCTTTTGAAGACATA 202402

QY 102 ACAGTTCACCTGAATGAATCTAAAGTCTTTGGTCTTTATCAGTGTGCTTC 153
Db 202401 ATACTTATGAGACCAATTCAGAAAGTCAATGCGATTTTCAGCAGGGTGAGTC 202350

RESULT 12
US-10-027-632-89184
; Sequence 89184, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89184
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-89184

Query Match 11.7%; Score 35.2; DB 13; Length 403;
Best Local Similarity 61.1%; Pred. No. 0.77;
Matches 55; Conservative 1; Mismatches 34; Indels 0; Gaps 0;

QY 39 RAATTTAACTCAGCCAGCCATTAACTGCTTAAAGTCTTAAAGT 128
Db 147 RAATCTCAAAATATGAGCTGACATTTGTCATGACCCAGATGATACAAATTTTGTAGGAG 206

QY 99 ACAACAGTTCACCTGAATGAATCTTAAAGT 128
Db 207 ACAACAGGAAAGCAATGAGGTATTTGAT 236

RESULT 13
US-10-027-632-89184
; Sequence 89184, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

Query Match 11.7%: Score 35.2; DB 13; Length 2286;

Search completed: August 4, 2005, 15:54:22  
Job time : 541.142 secs

---

*This Page Blank (uspto)*

FEATURES	Location/Qualifiers
source	1..300
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 300; DB 6; Length 300;
Best Local Similarity	100.0%; Pred. No. 2.4e-76;
Matches 300; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AAAATATAAAATTTAGCCAGGCTGATGTCATGTGCTCTGTAGTCCAGCTACTCTCGGAGG 60
Db	1 AAAATATAAAATTTAGCCAGGCGTGATGTGCTGTAGTCCAGCTACTCTCGGAGG 60
Qy	61 CTGAGCAGGAGAACCTCTTTGAATCCAGGAGGCGCAGGTTGCAGTGACGAGATAGTGC 120
Db	61 CTGAGGCGAGGAGAACCTCTTTGAATCCAGGAGGCGCAGGTTGCAGTGACGAGATAGTGC 120
Qy	121 CACTGCACCTCCAGCCTGGGTGACAGAGTAGTGAGCTGTCTCAAAAAATAAAATAAAATA 180
Db	121 CACTGCACCTCCAGCCTGGGTGACAGAGTAGTGAGCTGTCTCAAAAAATAAAATAAAATA 180
Qy	181 AAAAAATCGACACTGTGATTTCAGCAGGCTCTGGGTTGAAGCCCGAGAATCTCTCGATAAATTC 240

```

181 AAAAATGCAGACTGTAATCACAGGCTGGGTTGAGCCAGCACTCTGTGATAAATTC 240
241 AATGGCAGCTTAATCTACTTGGAGGTCATGGATGCCCTTGTCTAATCTAATAGAGCTACTGA 300
241 AATGGCAGCTTAATCTACTTGGAGGTCATGGATGCCCTTGTCTAATCTAATAGAGCTACTGA 300

RESULT 2
HSY18933 11793 bp DNA linear PRI 18-OCT-2000
LOCUS Homo sapiens MCP-1 gene and enhancer region.
DEFINITION Y18933
ACCESSION Y18933.1 GI:10933860
VERSION MCP-1 gene; monocyte chemoattractant protein-1.
KEYWORDS MCP-1 gene; monocyte chemoattractant protein-1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Finzer, P., Soto, U., Delius, H., Patzelt, A., Coy, J.F., Poustka, A.,
zur Hausen, H. and Roel, F.
TITLE Differential transcriptional regulation of the
monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and
non-tumorigenic HPV 18 positive cells: the role of the chromatin
structure and AP-1 composition
JOURNAL Oncogene 19 (29), 3235-3244 (2000)
MEDLINE 20374005
PUBMED 10918580
REFERENCE 2 (bases 1 to 11793)
AUTHORS Roel, F.
DIRECT SUBMISSION
TITLE Submitted (15-APR-1999) F. Roel, Applied Tumor Virology, DKPZ
JOURNAL (German Cancer Research Institute), Im Neuenheimer Feld 506, 69120
Heidelberg, FRG
FEATURES
source
1. .11793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
4906. .9073
/gene="MCP-1"
enhancer 4906. .4915
/gene="MCP-1"
enhancer 4935. .4944
/gene="MCP-1"
TATA_signal 7513. 7517
/gene="MCP-1"
mRNA join(<7550. .7685,8468. .8585,8968. .>9073)
/gene="MCP-1"
exon <7550. .7685
/gene="MCP-1"
number=1
CDS join(7610. .7685,8468. .8585,8968. .9073)
/gene="MCP-1"
/codon_start=1
/product="monocyte chemoattractant protein-1"
/protein_id="CAC14049.1"
/db_xref="GI:10933861"
/db_xref="GOA:P13500"
/db_xref="UniProt/Swiss-Prot:P13500"
/translation="MKVSNALLCLLLIATFIPQGLAOPDAINAVPTCCVNFNTNRKIS
VRLASVRRITSSKCPKEAVIFKTIVAKEICADPKQKWQDSMDHLDKQTQPTKT"
7686. .8467
/gene="MCP-1"
intron /number=1
exon 8468. .8585
/gene="MCP-1"
number=2
intron 8586. .8967
/gene="MCP-1"
number=2

```

```

exon 8968. .>9073
/gene="MCP-1"
/number=3
ORIGIN
Query Match 100.0%; Score 300; DB 9; Length 11793;
Best Local Similarity 100.0%; Pred. No. 2.7e-76; Indels 0; Gaps 0;
Matches 300; Conservative 0; Mismatches 0;
QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGTACTCCGGAGG 60
Db 2500 AAAATATAAAATAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGTACTCCGGAGG 2559
QY 61 CTGAGGAGGAGAACCTCTTGAATCCAGGAGCGCGAGTTGCAGTGAGCAGAGATAGTGC 120
Db 2560 CTGAGGAGGAGAACCTCTTGAATCCAGGAGCGCGAGTTGCAGTGAGCAGAGATAGTGC 2619
QY 121 CACTGCATCTCAGCGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 2620 CACTGCATCTCAGCGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 2679
QY 181 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAGCCAGCACTCTGTGATAAATTC 240
Db 2680 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAGCCAGCACTCTGTGATAAATTC 2739
QY 241 AATGCAGCTTAATCTACTTGGAGGTCATGGATGCCCTTGTCTAATCTAATAGAGCTACTGA 300
Db 2740 AATGCAGCTTAATCTACTTGGAGGTCATGGATGCCCTTGTCTAATCTAATAGAGCTACTGA 2799

RESULT 3
AC005549 147416 bp DNA linear PRI 22-SEP-1998
LOCUS Homo sapiens chromosome 17, clone hRPK.215_E13, complete sequence.
DEFINITION AC005549
ACCESSION AC005549.1 GI:3598724
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 147416)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone hRPK.215_E13
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 147416)
REFERENCE Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,
Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M.,
Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E.,
Devon, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W.,
Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G.,
Hagos, B., Heaford, A., Herena, L., Horton, L., Howland, J. C.,
Jactotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,
Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A.,
Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,
Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,
Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,
Ye, W. J., Zhao, J. and Zody, M.
DIRECT SUBMISSION
TITLE Submitted (27-AUG-1998) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 147416)
AUTHORS Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,
Boutwell, C., Brown, A., Castle, A., Cerny, J., Colangelo, M.,
Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E.,
Devon, K., Dewar, K., Donelan, L., Ferreira, P., Fitzhugh, W.,
Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G.,
Hagos, B., Heaford, A., Herena, L., Horton, L., Howland, J. C.,
Jactotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,
Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A.,
Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,
Subramanian, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,
Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,
Ye, W. J., Zhao, J. and Zody, M.

```

Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehocsky, J., Macdonald, P., Marquis, N., McSwan, P., McGuirk, A., McKernan, K., McNeil, J., Molla, M., Morris, M., Morrow, J., Mychalek, J. J., Nahr, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolsky, N., Toruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

**TITLE** Direct Submission

**JOURNAL**  
Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
**REFERENCE**  
4 (bases 1 to 147416)

## REFERENCE AUTHORS

**AUTHORS**

Birren, B., Fisman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Bunn, J., Boutwell, C., Brown, A., Castile, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Colliss, D., Depayre, E., Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W., Forreest, C., Funke, R., Gage, D., Gardyna, S., Geraisery, K., Grant, G., Hagos, B., Headford, A., Harena, L., Horton, L., Howland, J. C., Jacotot, I., Jones, C., Kann, L., Karatas, A., Lechocz, J., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Melndim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stielwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tsefaye, S., Tichovolsky, N., Toruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

**TITLE**

**JOURNAL**

COMMENT

On Sep 13, 1998 this sequence version replaced g1:3381743. All repeats were identified using RepeatMasker: Smit, A.F.A. &

All repeats were identified using RepeatMasker: 3M11C, A.F.A. & Green, P. (1996-1997)

## FEATURES

source

```
1. .147416
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
```

2 69 /CONE\_ID=KFC1-11 Human Bac Library

```
2.03
/rnt family="BIR1"
```

490  
/ EPC-1

450: .693  
/rnt fam!1v~n1.1ME3n

1001-1001

1291. .1390  
/vot famijr-umtpn

/rpt\_ram-3

2980. .3131  
/csm4 fcm4 j... mmmmmmm

```
/rpt_family="MER5A"
```

complement (3147).

/rpt\_family="MIR"

complement (4228).

```
/rpt_family="MLT11"
```

complement(5430. 5549)

```
/rpt_family="FLAM_C"
```

complément (5941).

/rpt fam:

6131. 6250

```
/rpt family="(TA)
```

6406. .6455

/rpt fam:

6574. .6600

0574. :0000  
/rpt family:

repeat_region	/rpt_family="AluJb"	6881. .6935
repeat_region	/rpt_family="MLT1D"	7119. .7360
repeat_region	/rpt_family="MLT1D"	7337. .7636
repeat_region	/rpt_family="MLT1D"	7689. .7988
repeat_region	/rpt_family="L1ME3"	8023. .8048
repeat_region	/rpt_family="AT_rich"	8871. .8996
repeat_region	/rpt_family="MIR"	10109. .10258
repeat_region	/rpt_family="MERSA"	complement(11031. .11072)
repeat_region	/rpt_family="MIR"	11567. .11836
repeat_region	/rpt_family="L2"	12501. .12981
repeat_region	/rpt_family="MER9"	complement(12983. .13257)
repeat_region	/rpt_family="AluJb"	13248. .15364
repeat_region	/rpt_family="MIR"	complement(16483. .16653)
repeat_region	/rpt_family="MIR"	16830. .16949
repeat_region	/rpt_family="MERSA"	17203. .17380
repeat_region	/rpt_family="LTR16A"	17393. .18306
repeat_region	/rpt_family="HERVL"	18315. .24307
repeat_region	/rpt_family="L1PA2"	24309. .24335
repeat_region	/rpt_family="AT_rich"	24513. .25245
repeat_region	/rpt_family="HERVL"	25613. .25758
repeat_region	/rpt_family="HERVL"	26135. .26228
repeat_region	/rpt_family="LTR16A"	complement(27814. .27950)
repeat_region	/rpt_family="MER9A"	28930. .28968
repeat_region	/rpt_family="AT_rich"	29960. .30341
repeat_region	/rpt_family="MLT1I"	32364. .32581
repeat_region	/rpt_family="MIR"	32601. .32902
repeat_region	/rpt_family="AluSx"	complement(33593. .39359)
repeat_region	/rpt_family="L1PA2"	39362. .39414
repeat_region	/rpt_family="A_rich"	complement(39920. .40047)
repeat_region	/rpt_family="MIR"	40112. .40151
repeat_region	/rpt_family="AT_rich"	complement(40152. .40442)
repeat_region	/rpt_family="AluJo"	complement(40810. .40898)
repeat_region	/rpt_family="MLT1J"	41293. .41469
repeat_region	/rpt_family="L1PA6"	complement(41846. .41956)
repeat_region	/rpt_family="puine-rich"	complement(42486. .42565)
repeat_region	/rpt_family="MLT1I"	42979. .43098
repeat_region	/rpt_family="MERSA"	

Tue Aug 9 17:01:27 2005

```
repeat_region 43554..43983
/rpt_family="LTR24"
repeat_region 44781..44811
/rpt_family="AT_rich"
repeat_region 46184..46296
/rpt_family="MIR"
repeat_region 47075..47140
complement(47075..47140)
/rpt_family="MLT1G"
repeat_region 47266..47307
complement(47266..47307)
/rpt_family="(CA)n"
repeat_region 47363..47442
complement(47363..47442)
/rpt_family="MLT1G"
repeat_region 48165..48209
complement(48165..48209)

Query Match 100.0%; Score 300; DB 9; Length 147416;
Best Local Similarity 100.0%; Pred. No. 2.9e-76;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGTCTGTAGTCCAGCTACTCGGAGG 60
Db 66527 AAAATATAAAATTAGCCAGGCGTGATGTCATGTCTGTAGTCCAGCTACTCGGAGG 66468
Qy 61 CTGAGGAGAGAACTCTTGAATCCAGAGGGCGAGGTTCAGTCCAGCAGAGATAGTC 120
Db 66467 CTGAGGAGAGAACTCTTGAATCCAGAGGGCGAGGTTCAGTCCAGCAGAGATAGTC 66408
Qy 121 CACTGCCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 66407 CACTGCCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 66348
Qy 181 AAAAATGCAGACTGTGATTCAGCAGGCTCTGGGTGAAGCCAGCAACTCTCTGATAATTC 240
Db 66347 AAAAATGCAGACTGTGATTCAGCAGGCTCTGGGTGAAGCCAGCAACTCTCTGATAATTC 66288
Qy 241 AATGGCACTTAACCTACTTGCAGTCTGAGTCTGGTCTTGTCTAATCTAATAGCACTACTCA 300
Db 66287 AATGGCACTTAACCTACTTGCAGTCTGAGTCTGGTCTTGTCTAATCTAATAGCACTACTCA 66228

RESULT 4
AX346554 5926 bp DNA linear PAT 01-FEB-2002
LOCUS Sequence 1625 from Patent WO020928.
DEFINITION AX346554
ACCESSION AX346554
VERSION AX346554.1 GI:18494440
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Olek,A.; Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 020928-A 1625 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
1..5926
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match 55.9%; Score 167.8; DB 6; Length 5926;
Best Local Similarity 80.7%; Pred. No. 6.4e-38;
Matches 196; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 58 AGGCTGAGGACAGAGAACTCTTGAATCCAGGAGCGGAGGTTCAGTCCAGCAGATAG 117
Db 1 AGGTTGAGTAGAGAGAAATTTTGAATTTAGAGCGGTAGGTTGTAGTGGTAGAGATAG 60
Qy 118 TGCCACTGCATCCAGCCTGGGTGACAGAGTGAAGTCTGTCTCAAAAAATAAATAA 177
Db 61 TGTATTGTATTATTAGTTTGGGTGATAGATGAGATTTTGTTTTAAAAATAAATAA 120
```

```
Qy 178 ATAAAAATGCAGACTGTGATTCAGCAGGCTCGGGTTGAAGCCAGCAACTCTCTGATAA 237
Db 121 ATAAAAATGCTAGATTGTGATTTAGTAGGTTGGGTTGAAGTTTGAATTTTGTATAA 180
Qy 238 TTCAATGGGCACTTAACCTACTTGGAGTCTATGATGCTTGTCTTAATCTAATAGAAGTAC 297
Db 181 TTTAATGGTATTTAATTAATTTGGAGGTATGATGTTTGTGTTAATTAATAGAAGTTAT 240
Qy 298 TGA 300
Db 241 TGA 243

RESULT 5
AC021222/c 177426 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 11 clone RP11-567M21, WORKING DRAFT
DEFINITION AC021222
ACCESSION AC021222.4 GI:7770692
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 177426)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 177426)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On May 12, 2000 this sequence version replaced gi:7233676.
```

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0567M21
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 154559 bases at least Q40
Consensus quality: 161804 bases at least Q30
Consensus quality: 166430 bases at least Q20
Insert size: 189000; agarose-fp
Insert size: 174526; sum-of-contigs
Quality coverage: 2.77 in Q20 bases; agarose-fp
Quality coverage: 3.06 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 1044: contig of 1044 bp in length
* 1045 1144: gap of unknown length
* 1145 2529: contig of 1385 bp in length
* 2530 2630: gap of unknown length
* 2630 3739: contig of 1110 bp in length
* 3740 3839: gap of unknown length
* 3840 5188: contig of 1349 bp in length
```

\* 5189 5288: gap of unknown length  
\* 5289 6861: contig of 1573 bp in length  
\* 6862 6961: gap of unknown length  
\* 6962 8387: contig of 1426 bp in length  
\* 8388 8487: gap of unknown length  
\* 8488 9957: contig of 1470 bp in length  
\* 9958 10057: gap of unknown length  
\* 10058 11934: contig of 1877 bp in length  
\* 11935 12034: gap of unknown length  
\* 12035 15639: contig of 3605 bp in length  
\* 15640 15739: gap of unknown length  
\* 15740 18805: contig of 3066 bp in length  
\* 18806 21750: gap of unknown length  
\* 18906 21751: contig of 2845 bp in length  
\* 21751 21850: gap of unknown length  
\* 21851 24673: contig of 2823 bp in length  
\* 24674 24773: gap of unknown length  
\* 24774 27874: contig of 3101 bp in length  
\* 27875 27974: gap of unknown length  
\* 27975 32401: contig of 4427 bp in length  
\* 32402 32501: gap of unknown length  
\* 32502 35859: contig of 3358 bp in length  
\* 35860 35959: gap of unknown length  
\* 35960 38608: contig of 2649 bp in length  
\* 38609 43983: gap of unknown length  
\* 43984 44084: gap of unknown length  
\* 44084 50370: contig of 6287 bp in length  
\* 50371 50470: gap of unknown length  
\* 50471 55419: contig of 4949 bp in length  
\* 55420 55519: gap of unknown length  
\* 55520 60146: contig of 4627 bp in length  
\* 60147 60246: gap of unknown length  
\* 60247 65990: contig of 5744 bp in length  
\* 65991 66091: gap of unknown length  
\* 66091 72519: contig of 6429 bp in length  
\* 72520 79368: contig of 6749 bp in length  
\* 79369 79468: gap of unknown length  
\* 79469 88631: contig of 9163 bp in length  
\* 88632 88731: gap of unknown length  
\* 88732 98171: contig of 9440 bp in length  
\* 98172 98271: gap of unknown length  
\* 98272 108874: contig of 10603 bp in length  
\* 108875 108974: gap of unknown length  
\* 108975 120856: contig of 11882 bp in length  
\* 120857 120957: gap of unknown length  
\* 120957 133833: contig of 12876 bp in length  
\* 133833 133932: gap of unknown length  
\* 133933 151337: contig of 17405 bp in length  
\* 151338 151437: gap of unknown length  
\* 151438 177426: contig of 25989 bp in length.

FEATURES  
source  
Location/Qualifiers  
1..177426  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="11"  
/clone="RP11-567M21"  
1..1044  
/note="assembly\_name:Contig10"  
1145..2529  
/note="assembly\_name:Contig12"  
2630..3739  
/note="assembly\_name:Contig13"  
3840..5188  
/note="assembly\_name:Contig15"  
5289..6861  
/note="assembly\_name:Contig16"  
6962..8387  
/note="assembly\_name:Contig17"  
8488..9957  
/note="assembly\_name:Contig18"

misc\_feature 10058..11934  
/note="assembly\_name:Contig19"  
misc\_feature 12035..15639  
/note="assembly\_name:Contig20"  
misc\_feature 15740..18805  
/note="assembly\_name:Contig21"  
misc\_feature 18906..21750  
/note="assembly\_name:Contig22"  
misc\_feature 21851..24673  
/note="assembly\_name:Contig23"  
misc\_feature 24774..27874  
/note="assembly\_name:Contig24"  
misc\_feature 27975..32401  
/note="assembly\_name:Contig25"  
misc\_feature 32502..35859  
/note="assembly\_name:Contig26"  
misc\_feature 35960..38608  
/note="assembly\_name:Contig27"  
misc\_feature 38709..43983  
/note="assembly\_name:Contig28"  
misc\_feature 44084..50370  
/note="assembly\_name:Contig29"  
misc\_feature 50471..55419  
/note="assembly\_name:Contig30"  
misc\_feature 55520..60146  
/note="assembly\_name:Contig31"  
misc\_feature 60247..65990  
/note="assembly\_name:Contig32"  
misc\_feature 66091..72519  
/note="assembly\_name:Contig33"  
misc\_feature 72620..79368  
/note="assembly\_name:Contig34  
clone\_end:77  
vector\_side:left"  
79469..88631  
/note="assembly\_name:Contig35"  
88732..98171  
/note="assembly\_name:Contig36"  
98272..108874  
/note="assembly\_name:Contig37"  
misc\_feature 108975..120856  
/note="assembly\_name:Contig38"  
misc\_feature 120957..133832  
/note="assembly\_name:Contig39"  
misc\_feature 133933..151337  
/note="assembly\_name:Contig40"  
misc\_feature 151438..177426  
/note="assembly\_name:Contig41"  
ORIGIN  
Query Match 54.7%; Score 164; DB 2; Length 177426;  
Best Local Similarity 84.1%; Pred.No. 9e-37; Indels 0; Gaps 0;  
Matches 185; Conservative 0; Mismatches 35;  
QY 1 AAAATATATAAATTAGCCAGGCGTGATGTGCTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
Db 100140 AAAATACAAAATTAGCCAGGCGTGATGTGCTCATGTGCTGTAGTCCAGCTACTCGGAGG 100081  
QY 61 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGCGGAGGTTGCGATGAGCAGAGATAGTGC 120  
Db 100080 CTGAGAGAGGAGAAATTTGCTTGAACCCAGGAGGAGGTTGCGATGAGCAGATCGCAC 100021  
QY 121 CACTGCATCTCCAGCTCGGTGACAGATGAGAGTCTGTCTCAAAAAATAAAATAAATA 180  
Db 100020 CACTGCATCTCCAGCTCGGTGACAGATGAGAGTCTGTCTCAAAAAATAAAATAAATA 99961  
QY 181 AAAAATGCAGACTGTGATTTCAGCAGGTTCTGGGTGAAGCC 220  
Db 99960 AAAAATAAAGAAATGCTAAGCAACAGCCAGCGCATGGTGC 99921  
RESULT 6  
AP003041/c

LOCUS AP003041 188172 bp DNA linear PRI 10-AUG-2002  
DEFINITION Homo sapiens genomic DNA, chromosome 11 clone:RP11-567M21, complete sequence.  
ACCESSION AP003041  
VERSION AP003041.3 GI:22202826  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
AUTHORS Published Only in Database (2000)  
TITLE Homo sapiens genomic DNA  
JOURNAL  
REFERENCE Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
AUTHORS Direct Submission  
TITLE Submitted (11-DEC-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:natori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)  
JOURNAL  
COMMENT On Aug 9, 2002 this sequence version replaced gi:20334333.  
FEATURES  
source  
1. .188172  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q"  
/clone="RP11-567M21"  
ORIGIN  
Query Match 54.7%; Score 164; DB 9; Length 188172;  
Best Local Similarity 84.1%; Pred. No. 9e-37;  
Matches 185; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 AAAAAATATAAAATAGCAGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
Db 78956 AAAAAATATAAAATAGCAGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
QY 61 CTGAGGAGGAGAACTCTTGAATCCAGGAGGCGAGGTTGCAGTGCAGGAGATAGTGC 120  
Db 78896 CTGAGGAGGAGAACTCTTGAATCCAGGAGGCGAGGTTGCAGTGCAGGAGATAGTGC 120  
QY 121 CACTGCACCTCCAGCTCGGTGACAGAGTGAGACTCTGTCTCAAAAAATATAAAATA 180  
Db 78836 CACTGCACCTCCAGCTCGGTGACAGAGTGAGACTCTGTCTCAAAAAATATAAAATA 180  
QY 181 AAAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAAGCC 220  
Db 78776 AAAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAAGCC 220  
RESULT 7 AC010290 238472 bp DNA linear PRI 23-OCT-2001  
LOCUS Homo sapiens chromosome 5 clone CTB-125B20, complete sequence.  
DEFINITION AC010290  
ACCESSION AC010290  
VERSION AC010290.7 GI:16328257  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 238472)  
TITLE Direct Submission  
JOURNAL  
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 238472)  
AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission  
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 238472)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (23-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Oct 23, 2001 this sequence version replaced gi:7711412.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.7% of Sequence;  
Estimated Total Number of Errors is 0.5.  
FEATURES  
source  
1. .238472  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTB-125B20"  
ORIGIN  
Query Match 54.7%; Score 164; DB 9; Length 238472;  
Best Local Similarity 84.1%; Pred. No. 9.1e-37;  
Matches 185; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 AAAATATAAAATAGCAGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
Db 171312 AAAATATAAAATAGCAGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
QY 61 CTGAGGAGGAGAACTCTTGAATCCAGGAGGCGAGGTTGCAGTGCAGGAGATAGTGC 120  
Db 171252 CTGAGGAGGAGAACTCTTGAATCCAGGAGGCGAGGTTGCAGTGCAGGAGATAGTGC 120  
QY 121 CACTGCACCTCCAGCTCGGTGACAGAGTGAGACTCTGTCTCAAAAAATATAAAATA 180  
Db 171192 CACTGCACCTCCAGCTCGGTGACAGAGTGAGACTCTGTCTCAAAAAATATAAAATA 180  
QY 181 AAAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAAGCC 220  
Db 171132 AAAAAATGCAGACTGTGATTCAGCAGGTCCTGGTTGAAGCC 220  
RESULT 8 HS28H20/c 127418 bp DNA linear PRI 19-APR-2001  
LOCUS Human DNA sequence from clone Rp1-28H20 on chromosome 20q13.1  
DEFINITION Contains the SLC2A10 gene encoding a solute carrier family 2 (facilitated glucose transporter) member 10, the 5' end of a novel gene, ESTs, STSs, GSSs and three CpG islands, complete sequence.  
ACCESSION AL031055  
VERSION AL031055.1 GI:4375937  
KEYWORDS HTG; CpG island; SLC2A10.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 127418)  
TITLE Direct Submission  
JOURNAL Submitted (18-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hunquery@sanger.ac.uk requests: clonerequest@sanger.ac.uk  
COMMENT On Mar 7, 1999 this sequence version replaced gi:4056528.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:  
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TRMBL; Wp:, WORMPEP; Information  
 on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 20, constructed by the Sanger Centre Chromosome 20  
 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>  
 This sequence is the entire insert of clone Rpl-28H20. This sequence  
 was finished as follows unless otherwise noted: all regions were  
 either double-stranded or sequenced with an alternate chemistry or  
 covered by high quality data (i.e., phred quality >= 30); an  
 attempt was made to resolve all sequencing problems, such as  
 compressions and repeats; all regions were covered by at least one  
 plasmid subclone or more than one M13 subclone; and the assembly  
 was confirmed by restriction digest. Rpl-28H20 is from the library  
 RPCI-1 constructed by the group of Pieter de Jong. For further  
 details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2.

## FEATURES

## source

## Location/Qualifiers

1..127418  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="q13.1"  
 /clone="Rpl-28H20"  
 /clone\_lib="RPCI-1"  
 complement(1..3110)  
 /genes="dJ101A2.2"  
 complement(join(<1..562,2463..3110))  
 /genes="dJ101A2.2"  
 /product="dJ28H20.2 (novel protein)"  
 /notes="match: CDNAS: Em:AK013049 Em:AK011574  
 match: ESTs: Em:AA233726 Em:T87841 Em:AA756018 Em:AA113277  
 Em:AA985248 Em:AA682750 Em:AA461487 Em:A1028598  
 Em:AA659710 Em:A1149891 Em:AA579641 Em:AA516077  
 Em:AA460132 Em:AA483490 Em:AA046575"  
 /evidence="not\_experimental"  
 18..494  
 /notes="match: STS: Em:HS28H20T"  
 complement(join(84..562,2463..2745))  
 /genes="dJ101A2.2"  
 /notes="supported by FGESH and GENSCAN  
 continues in Em:AL133520 as dJ101A2.2"  
 /codon\_start=1  
 /evidence="not\_experimental"  
 /product="dJ28H20.2 (novel protein)"  
 /protein\_id="CAC00561.1"  
 /db\_xref="GI:9588402"  
 /db\_xref="GOA:Q96S44"  
 /db\_xref="UniProt/Swiss-Prot:Q96S44"  
 /translation="MAARATTPADGEEPAEALAAARERSRFLGLELVKQAAE  
 AVFGRFGRAAVIKRPFGYHPALRGRTVQEARLLRCRAGISAPVVE  
 FVDYASNCILMEIEGVSIVRYDIQSTMETETPQGLSLAKTIQVILARMHDEDLIH  
 GDLTTSNMLKPLPLENLVLDGLSFTSALPEDKGVLDVLEKAFSLSTHENTETVF  
 EAFLSKYSTSSKKARVPLKLDVLRGRKRMVG"  
 867..1145  
 /note="AluX repeat: matches 1..298 of consensus"  
 1251..1353  
 /notes="MIR repeat: matches 50..151 of consensus"  
 1672..1797  
 /note="MIR repeat: matches 8..135 of consensus"  
 1855..2079  
 /notes="MIR repeat: matches 33..262 of consensus"  
 2155..2248  
 /notes="MIR repeat: matches 33..135 of consensus"  
 2384..3150  
 /note="CpG island"  
 /evidence="not\_experimental"  
 2812..2871  
 /note="3 copies 20 mer 85% conserved"

misc\_feature  
 complement(2876..3110)  
 /gene="dJ101A2.2"  
 /note="match: STS: Em:G15621"  
 3865..3962  
 /note="MER21B repeat: matches 299..391 of consensus"  
 3963..5655  
 /note="L1MA8 repeat: matches 4532..6270 of consensus"  
 5669..5895  
 /note="L1ME repeat: matches 5501..5739 of consensus"  
 6332..6548  
 /note="L2 repeat: matches 2524..2746 of consensus"  
 6626..6915  
 /note="AluX repeat: matches 3..290 of consensus"  
 6918..6980  
 /note="MER94 repeat: matches 1..62 of consensus"  
 7023..7322  
 /note="AluDo repeat: matches 1..312 of consensus"  
 7323..7417  
 /note="L2 repeat: matches 2389..2513 of consensus"  
 7457..7623  
 /note="MIR repeat: matches 93..262 of consensus"  
 7864..7934  
 /note="MLTII repeat: matches 311..382 of consensus"  
 8046..8240  
 /note="MLTIJ repeat: matches 9..211 of consensus"  
 8347..8647  
 /note="AluSq repeat: matches 1..300 of consensus"  
 8685..8793  
 /note="MLTIJ repeat: matches 271..366 of consensus"  
 9095..9221  
 /note="MIR repeat: matches 102..242 of consensus"  
 9647..9776  
 /note="L2 repeat: matches 2170..2303 of consensus"  
 9793..10028  
 /note="AluDo repeat: matches 74..306 of consensus"  
 10038..10206  
 /note="LTR16C repeat: matches 126..316 of consensus"  
 10631..10690  
 /note="MLTID repeat: matches 445..505 of consensus"  
 10691..10995  
 /note="AluDb repeat: matches 1..295 of consensus"  
 10996..11200  
 /note="MLTID repeat: matches 248..445 of consensus"  
 11232..11634  
 /note="MSTB repeat: matches 2..409 of consensus"  
 11670..11761  
 /note="MLTID repeat: matches 128..192 of consensus"  
 11762..12062  
 /note="AluX repeat: matches 1..293 of consensus"  
 12063..12180  
 /note="MLTID repeat: matches 1..128 of consensus"  
 complement(12181..12344)  
 /note="match: GSS: Em:AQ190762"  
 13011..13129  
 /note="MLTIJ repeat: matches 252..368 of consensus"  
 13210..13392  
 /note="MER20 repeat: matches 1..215 of consensus"  
 13397..13442  
 /note="MIR repeat: matches 104..146 of consensus"  
 13660..13783  
 /note="LTR33 repeat: matches 330..452 of consensus"  
 13894..14254  
 /note="L2 repeat: matches 2255..2619 of consensus"  
 14305..14435  
 /note="L2 repeat: matches 2617..2748 of consensus"  
 14446..14525  
 /note="4 copies 20 mer 76% conserved"  
 15476..15752  
 /note="AluJb repeat: matches 2..278 of consensus"  
 15774..15952  
 /note="MERS8A repeat: matches 42..220 of consensus"  
 16190..16304  
 /note="L2 repeat: matches 2657..2744 of consensus"



```
repeat_region /rpt_family="MLT1"
complement(10611..10904)
repeat_region /rpt_family="Alu"
11160..11445
repeat_region /rpt_family="Alu"
complement(11721..11961)
CDS /rpt_family="Alu"
join(12984..13098,13822..13945,14905..14990,15085..15206,
15568..15755,15846..15986,17475..17550,17865..17941,
18124..18221,19545..19619,21008..21096,21207..21278,
21370..21558)
/notes="Hypothetical 56kDa human ATP-dependent RNA
helicase; Putative ATP-dependent RNA helicase of DEAD box
family. Most similar (57% identical) to probable
ATP-dependent RNA helicase Ddp45A (S38329) - fruit fly
(Drosophila melanogaster)
Human EST matches: AA534472, H08289, AA464032, AA196836,
AA569862, AA573466, AA464741, W46150, AA535538, AA488261,
W46162, R00974, AA378518, R26194, R15246, AA359638, etc.
Mouse EST matches: AA271508
Drosophila EST matches: AA540379
Rat EST match: H35240"
/codon_start=1
/product="R27090_2"
/protein_id="AAB81544.1"
/db_xref="GI:2443870"
/translation="MAGFAELGLSSMLVQCRQLGKOPTPVQLGCI PAILEGRDCIG
LIIVGSGKTAQVFLPILOKLSDEPGYIGFCLVLTPTRELAVQIAEOPVLGKPLGLKDC
LIIVGMDVAAQALESRKPHVATPGRLADHLRSNTFSIKIRFLVNDLRLLEQ
GCTDFTVLEAILAAPARQTLTFSATLDTLRELOGLATNPFPWEAQAPSVTEQ
LDQRYLLPEKVKADYVHLIRFQDEHEDWSIIIFNTCKTCQILCMMLRFPSPVY
ALHSMKQKEFAALAKFKSSIVRIIADTASRGIDIPTVQVINHNTPGPKPIYH
RVGRTAAGRGQAIPLVQYDIHLVHAIEBOIKKLEBFSVEAEVLQILQVNVVR
RECEIKLEAHFDEKEINKRKQILLEGKDPLEAKKAEKELAKIKQNRPFKEVET
LRQKAGRAGHGRPRPTSGSHGVPVPSQGLV"
14112..14397
/rpt_family="Alu"
16503..16734
repeat_region /rpt_family="Alu"
complement(17052..17340)
repeat_region /rpt_family="Alu"
18380..18802
repeat_region /rpt_family="Alu"
20044..20821
misc_feature /rpt_family="Alu"
complement(21577..21650)
/notes="BLASTN similarity to Z41188 (216..289); match:
0.85, score: 3.5e-33; database searched: est; H. sapiens
partial cDNA sequence"
/misc_feature /notes="BLASTN similarity to T16014 (1..261); match: 0.94,
score: 1.1e-90; database searched: est; IB249 Infant
brain, Bento Soares Homo sapiens cDNA 3'end."
/misc_feature /notes="BLASTN similarity to Z41188 (42..252); match: 0.92,
score: 8.5e-89; database searched: est; H. sapiens partial
cDNA sequence"
misc_feature /complement(21799..21865)
/notes="BLASTN similarity to Z41188 (1..67); match: 0.94,
score: 8.5e-89; database searched: est; H. sapiens partial
cDNA sequence"
CDS complement(join(22685..22876,24567..24653,24751..24867,
25218..25374,26166..26287,27506..27652,31595..31726,
31970..32136,32211..32224))
/notes="Hypothetical 41.3 kDa human protein most similar to
Vesl and GLGF proteins of rat; Residues 1-128 of
hypothetical protein R27090_3 are 81% identical to Rat
GLGF protein (U92079) and 83% identical to Rat Vesl
protein (AB003726); both proteins are expressed in brain
and upregulated during seizures.
Human EST matches: W42820, AA121538, W42730, AA127702
Mouse EST matches: AA407944, AA408331, AA013888, AA035853,
AA212542
```

```
Drosophila EST matches: AA391781, AA202338, AA201147,
AA202832, AA246370"
/codon_start=1
/product="R27090_3"
/protein_id="AAB81545.1"
/db_xref="GI:2443871"
/translation="MSTAREQPIFSTRAHFQIDPATKRWIPAGKHALTVSYFYDAT
RNVYRIISIGGAKAIINSTVTQPMFTTKTSQKFGOWADSRANTVYGLGFASEOHLTOP
ASKFOVEKAAARLAREKSDGGELSPALGLASHOVSTYSPMPAWAPVPPSPLVAN
GPGEKLPKRSQADAPGPTERERLAKMLSEGSGVEQWAEFFALQDSNNKLALRE
ANAAQRQQLQEAERLURQKRVAELEAAQAASEVPTGKEGIGGQSGQLEQLEAL
VTKDQETQLKSQTGGPREALEAARESTQOKVQDLETRNAELHQLRAMERLSLEA
RAERARAEGVRAAQLLDVSLFSELSEGLRLAEEAP"
complement(23068..23216)
/rpt_family="Alu"
repeat_region 23231..23507
repeat_region /rpt_family="Alu"
23533..24266
repeat_region /rpt_family="Alu"
complement(26954..27247)
repeat_region /rpt_family="Alu"
complement(27925..28084)
repeat_region /rpt_family="Alu"
complement(28109..28363)
repeat_region /rpt_family="Alu"
complement(28377..28665)
repeat_region 28980..29921
repeat_region /rpt_family="Alu"
29950..30567
repeat_region /rpt_family="Alu"
30657..30931
/rpt_family="Alu"
Query Match 53.6%; Score 160.8; DB 9; Length 38041;
Best Local Similarity 86.8%; Pred. No. 7.3e-36;
Matches 177; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY 1 AAAAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCCGGAGG 60
|||||
Db 27125 AAAAAATACAAAATTAGCCAGGCGTGTCGTCATGCGCTGTAGTCCAGCTACTCCGGAGG 27066
QY 61 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGCGGAGTTGTCAGTGAGCAGAGATAGTGC 120
|||||
Db 27065 CTGAGGCGAGGAGAAATCGCTTGAACCCAGGAGATGGAGTTGTCAGTGAGCAGATCGTGC 27006
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAATAAATA 180
|||||
Db 27005 CACTGCACCTCCAGCCTGGGTGACAGAGTGTGACTCTGTCTCAAAAAATAAATAAATA 26946
QY 181 AAAAAATCGAGACTGTGATTTCAGCA 204
Db 26945 GAATGGCTTAAGTGTGAGCAGAA 26922
```

```
RESULT 10
LOCUS AC012094/c
DEFINITION Homo sapiens chromosome 3 clone RP11-10315, *** SEQUENCING IN
PROGRESS ***, 21 unordered pieces.
ACCESSION AC012094 AC012029
VERSION AC012094.16 GI:20279325
KEYWORDS HTG; HTGS PHASE1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 80952)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oeman,F.R., Allen,C.,
Albrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhaey,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
```

\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* As soon as it is available and the accession number will  
\* be preserved.

preserved.	1	2359:	contig	of 2359	bp in length
	2360	2459:	gap	of unknown	length
	2460	4662:	contig	of 2203	bp in length
	4663	4762:	gap	of unknown	length
	4763	7217:	contig	of 2455	bp in length
	7218	7317:	gap	of unknown	length
	7318	10240:	contig	of 2323	bp in length
	10241	10340:	gap	of unknown	length
	10341	12454:	contig	of 2114	bp in length
	12455	12554:	gap	of unknown	length
	12555	15441:	contig	of 2887	bp in length
	15442	15441:	gap	of unknown	length
	15443	18194:	contig	of 2653	bp in length
	18195	18294:	gap	of unknown	length
	18295	20599:	contig	of 2305	bp in length
	20600	20699:	gap	of unknown	length
	20700	24735:	contig	of 4036	bp in length
	24736	24835:	gap	of unknown	length
	24836	27247:	contig	of 2412	bp in length
	27248	27347:	gap	of unknown	length
	27348	30324:	contig	of 2977	bp in length
	30325	30424:	gap	of unknown	length
	30425	3317:	contig	of 3293	bp in length
	33178	33817:	gap	of unknown	length
	33818	36374:	contig	of 2557	bp in length
	33819	36474:	gap	of unknown	length
	36375	40161:	contig	of 3687	bp in length
	36475	40261:	gap	of unknown	length
	40162	40261:	gap	of unknown	length
	40262	45330:	contig	of 5069	bp in length
	40263	45330:	gap	of unknown	length
	45331	49718:	contig	of 4288	bp in length
	45332	49818:	gap	of unknown	length
	49719	56287:	contig	of 6469	bp in length
	49819	56387:	gap	of unknown	length
	56388	60054:	contig	of 3667	bp in length
	56389	60054:	gap	of unknown	length
	60055	60154:	contig	of 5473	bp in length
	60155	65627:	gap	of unknown	length
	65628	73013:	contig	of 7286	bp in length
	65728	73014:	gap	of unknown	length
	73014	90829:	contig	of 7819	bp in length

```
Location/Qualifiers
1. .80952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="pD11-10315"
```

## ORIGIN

Query Match 53.3%; Score 160; DB 2; Length 80952;  
Best Local Similarity 89.6%; Pred. No. 1.3e-35;  
Matches 172; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY	1	AAATATATAAATTAAGCAGGCGTGCATGTCATGTGCTTGTAGTCCAGCTACTCGGAGG	60
Db	47930	AAATATATAAATTAAGCAGGCGTGTGGCATGTGCTTGTATCCAGCTACTCGGAGG	47871
QY	61	CTGAGGCGAGGAGAACTCTTGAATCCAGAGGCGCAGGTCCTCAGTGACGAGATAGTGC	120
Db	47870	CTGAGGCGAGGAGAACTCGTTGAACCCAGAGGCGGAGGTTGCATGTGCGCGAGTCATGC	47811
QY	121	CACGTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAATAAATA	180
Db	47810	CACGTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAATAAATAAATAAATA	47751

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently



QY	121	CACTGCACCTCCAGCCCTGGGTGACAGAGTGGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db	96142	CATGCACTCCAGCCCTGGGTGACAGAGTGGAGACTCTGTCTCAAAAAATAAAATAAATA 96083
QY	181	AAAAATGCAGAC 192
Db	96082	AAATATATACAC 96071
RESULT 12		
AL445685/c		
LOCUS	142790 bp	DNA linear PRI 21-APR-2002
DEFINITION	Human DNA sequence from clone RP11-91A18 on chromosome 1, complete sequence.	
ACCESSION	AL445685	
VERSION	AL445685.18	GI:20302179
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1 Wallis, J.	
AUTHORS	Direct Submission	
TITLE	Submitted (20-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk	
JOURNAL	On Apr 24, 2002 this sequence version replaced gi:14330007. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	
COMMENT	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1 RP11-91A18 is from the library RPC1-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBac3.6.	
FEATURES	Location/Qualifiers	
source	1..142790	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="1"	
	/clone="RP11-91A18"	
	/clone_lib="RPC1-11.1"	
ORIGIN		
Query Match	53.1%;	Score 159.4; DB 9; Length 142790;
Best Local Similarity	77.5%;	Pred. No. 1.9e-35;
Matches 193; Conservative	0;	Mismatches 56; Indels 0; Gaps 0;
QY	1	AAAAATAAAAAATTAGCCAGGCGTGATGTCATGTCCTCTAGTCCAGCTACTCGGAGG 60
Db	129069	AAAAACAAAAAATTAGCCAGGCGTGATGTCCTCTAGTCCAGCTACTCGGAGG 129010

QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCAGGTTGCACTGAGCAGAGATAGTGC 120  
 |||||  
 Db 129009 CTGAGCAGGAGAACTCGTTGAACCCAGGAGGTGGAGTTGCACTGAGCGAGATCACGC 128950  
 |||||  
 QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAAAAATA 180  
 |||||  
 Db 128949 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAAAAATA 128890  
 |||||  
 QY 181 AAAAATGCAGACTGTGATTCAGCAGGTCTGGGTGAAGCCAGCAACTCTCTGATAAATTC 240  
 |||||  
 Db 128889 AAAGCGGAGCGCGGCATCAGCAGGTGAGGATCAAGCAAGACCATCTCTGCTAAAT 128830  
 |||||  
 QY 241 AATGGCACT 249  
 |||||  
 Db 128829 GGTGAACT 128821

RESULT 13  
 AC146127 190149 bp DNA linear PRI 09-JUN-2004  
 DEFINITION Pan troglodytes BAC clone RP43-23F10 from 7, complete sequence.  
 AC146127  
 ACCESSION  
 VERSION  
 AC146127.4 GI:47903402  
 KEYWORDS  
 HTG.

ORGANISM  
 SOURCE Pan troglodytes (chimpanzee)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE  
 AUTHORS Trani, L., Haglund, K. and Meyer, R.  
 TITLE The sequence of Pan troglodytes BAC clone RP43-23F10  
 JOURNAL Unpublished (2001)  
 REFERENCE 2 (bases 1 to 190149)  
 AUTHORS Wilson, R.K.

TITLE Direct Submission  
 JOURNAL Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 190149)  
 AUTHORS Wilson, R.K.

TITLE Direct Submission  
 JOURNAL Submitted (27-APR-2004) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 4 (bases 1 to 190149)  
 AUTHORS Wilson, R.K.

TITLE Direct Submission  
 JOURNAL Submitted (02-JUN-2004) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 5 (bases 1 to 190149)  
 AUTHORS Wilson, R.K.

TITLE Direct Submission  
 JOURNAL Submitted (09-JUN-2004) Washington University School of Medicine,  
 Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO  
 63108, USA

COMMENT On Jun 2, 2004 this sequence version replaced gi:46576149.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@wustl.wustl.edu  
 ----- Summary Statistics  
 -----  
 Center project name: C\_PT023F10

## NOTICE:

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. Wes Warren,  
 Department of Genetics, Washington University, St. Louis MO. For  
 additional information about the map position of this sequence, see  
 http://genome.wustl.edu

SOURCE INFORMATION:  
 The RPCI-43 BAC library has been constructed by Chung-Li Shu. DNA  
 was isolated from white blood cells obtained from a male chimpanzee  
 (Pan troglodytes, 'Clint', Yerkes #C0471; birthdate: 6-6-80). The  
 clone and detailed information can be obtained from ResGen  
 (http://www.resgen.com) or Pieter de Jong and co-workers at  
 http://www.bacpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:  
 This sequence is the entire insert of the clone.

## FEATURES

## source

Location/Qualifiers  
 1..190149  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /chromosome="7"  
 /map="7"  
 /clone="RP43-23F10"  
 /clone\_lib="RPCI-43"  
 128925..128976  
 /note="Sequence derived from one plasmid subclone."  
 128996..129012  
 /note="Sequence derived from one plasmid subclone."

unsure

unsure

## ORIGIN

Query Match 52.6%; Score 157.8; DB 9; Length 190149;  
 Best Local Similarity 90.8%; Pred. No. 5.7e-35;  
 Matches 168; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 AAAATATATAAATTAGCCAGCGGTGATGTCATGTGCTGTAGTCCAGTCTCTGAGGAGG 60  
 |||||  
 Db 70073 AAAATACAAAAATTAGCCCGGTGTGTGTCATGTGCTGTAATCCAGCTACTCTCGGAGG 70132  
 |||||  
 QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCAGGTTGCACTGAGCAGAGATAGTGC 120  
 |||||  
 Db 70133 CTGAGCAGGAGAACTCGTTGAACCCAGGAGCGGAGTTGCACTGAGCGAGATCATGC 70192  
 |||||  
 QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAAAAATA 180  
 |||||  
 Db 70193 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAAAAATA 70252  
 |||||  
 QY 181 AAAA 185  
 |||||  
 Db 70253 AAAA 70257

## RESULT 14

## AC146483

## LOCUS

## DEFINITION

## AC146483

## ACCESSION

## VERSION

## AC146483.2

## GI:37515257

## KEYWORDS

## HTG.

## SOURCE

## Pan troglodytes (chimpanzee)

## ORGANISM

## Pan troglodytes

## Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

## REFERENCE

## 1 (bases 1 to 208066)

## AUTHORS

## Trani, L., Cotton, M., Haglund, K. and Bielicki, L.

## TITLE

## The sequence of Pan troglodytes BAC clone RP43-30K12

## JOURNAL

## Unpublished (2001)

## REFERENCE

## 2 (bases 1 to 208066)

## AUTHORS

## Sulston, J.E. and Wilson, R.

## TITLE

## Sequencing of Pan troglodytes

## JOURNAL

## Unpublished (2001)

## REFERENCE

## 3 (bases 1 to 208066)

## AUTHORS

## Wilson, R.K.

61 CTGAGCAGAGAACTCTTGAATCCAGAGCGGAGTTGCGTGCAGTGCAGATAGTGC 120  
 134813 CTGAGCAGAGAACTCTTGAATCCAGAGCGGAGTTGCGTGCAGTGCAGATAGTGC 134872  
 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATA 180  
 134873 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATA 134932  
 181 AAAAA 185  
 134933 AATA 134937

RESULT 15  
 AC146006/c  
 LOCUS  
 DEFINITION Pan troglodytes BAC clone RP43-165A19 from 7, complete sequence.  
 ACCESSION AC146006  
 VERSION AC146006.2 GI:40286578  
 KEYWORDS HTG.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 REFERENCE  
 AUTHORS Trani L., Cotton M. and Haakenson W.  
 TITLE The sequence of Pan troglodytes BAC clone RP43-165A19  
 JOURNAL Unpublished (2001)  
 REFERENCE  
 AUTHORS Wilson R.  
 TITLE Sequencing of Pan troglodytes  
 JOURNAL Unpublished (2001)  
 REFERENCE  
 AUTHORS Wilson R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE  
 AUTHORS Wilson R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-AUG-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE  
 AUTHORS Wilson R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JAN-2004) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Dec 21, 2003 this sequence version replaced gi:33986957.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@wustl.edu  
 ----- Summary Statistics  
 Center project name: C\_PT165A19  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Direct Submission  
 Submitted (19-AUG-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 4 (bases 1 to 208066)  
 Wilson R.K.  
 Direct Submission  
 Submitted (04-OCT-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 5 (bases 1 to 208066)  
 Wilson R.  
 Direct Submission  
 Submitted (29-OCT-2003) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Oct 4, 2003 this sequence version replaced gi:33859971.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@wustl.edu  
 ----- Summary Statistics  
 Center project name: C\_PT030K12  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION:  
 The RPCI-43 BAC Library has been constructed by Chung-li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Cint', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (http://www.resgen.com) or Pieter de Jong and co-workers at http://www.bacpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:  
 This sequence is the entire insert of the clone.

FEATURES  
 source  
 1.208066  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /chromosome="7"  
 /map="7"  
 /clone="RP43-30K12"  
 /clone\_lib="RPCI-43"  
 1..349  
 /note="Sequence derived from one plasmid subclone."

ORIGIN  
 Query Match 52.6%; Score 157.8; DB 9; Length 208066;  
 Best Local Similarity 90.8%; Pred. No. 5.7e-35;  
 Matches 168; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 1 AAAAAATAAAATAGCCAGGCGTGTATGTCATGTCCTAGTCCAGCTACTCGGAGG 60  
 134753 AAAAAATAAAATAGCCAGGCGTGTATGTCATGTCCTAGTCCAGCTACTCGGAGG 134812

restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clinton', Yerkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC146236.

#### FEATURES

```

source
1. .211780
   /organism="Pan troglodytes"
   /mol_type="genomic DNA"
   /db_xref="taxon:9598"
   /chromosome="7"
   /map="7"
   /clone="RP43-165A19"
   /clone_lib="RPCI-43"
   /11047..111070
   /note="Sequence derived from one plasmid subclone."
unsure
111090..111244
   /note="Sequence derived from one plasmid subclone."
unsure
143326..143405
   /note="Sequence derived from PCR product of project DNA."
unsure
160484..160546
   /note="Sequence derived from one plasmid subclone."

```

#### ORIGIN

```

Query Match          52.6%; Score 157.8; DB 9; Length 211780;
Best Local Similarity 90.8%; Pred. No. 5.7e-35;
Matches 168; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 169993 AAAATACAAAATTAGCCCGGTGTGGTGGCATGTGCTGTATCCAGCTACTCGGGAGG 169934

QY 61 CTGAGCGAGGAGACCTCTTGATCCAGGAGCGGAGGTTCAGTCAGCAGAGATAGTGC 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 169933 CTGAGCGAGGAGAAATCGCTTGAACCCAGGAGCGGAGGTTCAGTCAGTCAGATCATGC 169874

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 169873 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAATAAATAAATAA 169814

QY 181 AAAAA 185
    |||||
Db 169813 AAATA 169809

```

Search completed: August 4, 2005, 11:17:00  
Job time : 1211.83 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 3, 2005, 15:09:47 ; Search time 198.826 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-6  
Perfect score: 300  
Sequence: 1 aaatataaaattagccag.....aatctaataagaactactga 300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	6	ADH13943
2	300	100.0	11793	12	AD003803
3	167.8	55.9	5926	6	ABL33652 Human Ccl
4	160.8	53.6	12026	4	AAK67211 Human imm
5	160.8	53.6	31926	13	ABD33528 Human can
6	156.4	52.1	5304	4	ABK42328 Genomic s
7	156.4	52.1	5304	4	ABK42327 Genomic s
8	156.4	52.1	5304	9	ADB60483 Connectiv
9	156.4	52.1	5304	9	ADB60484 Connectiv
10	156.4	52.1	9796	6	ABK69846 Human sec
11	156.4	52.1	9796	8	ACC50861 Human sec
12	156.4	52.1	9796	8	ABZ71483 Secreted
13	156.4	52.1	9796	9	ADB91839 Human sec
14	156.4	52.1	9796	10	ADC74614 Human sec
15	156.4	52.1	9796	10	ADD38134 cDNA clon
16	156.4	52.1	9796	10	ADA57728 BAC fragm
17	156.4	52.1	20245	6	AAK72318 Human imm
18	156.4	52.1	20245	6	ABK69845 Human sec
19	156.4	52.1	20245	8	ACC50860 Human sec
20	156.4	52.1	20245	8	ABZ71482 Secreted

21	156.4	52.1	20245	9	ADB91838 Human sec
22	156.4	52.1	20245	10	ADC74613 Human sec
23	156.4	52.1	20245	10	ADD38133 cDNA clon
24	156.4	52.1	20245	10	ADA57727 BAC fragm
c 25	155.8	51.9	195917	12	ADQ20606 Human sof
26	155.4	51.8	252907	13	ABD33694 Human can
27	154.6	51.5	5021	4	AAK81320 Human imm
c 28	153.8	51.3	118584	10	ADC87623 Human GPC
c 29	153.8	51.3	349999	10	AD87010 Human GPC
c 30	153.6	51.2	232	8	ABZ74462 Secreted
c 31	153.6	51.2	232	8	ADA98791 Secreted
c 32	153.6	51.2	232	10	ABZ67828 Human sec
c 33	153	51.0	144723	11	ACN44898 Human gen
c 34	152.8	50.9	83517	13	ABD32581 Human can
c 35	152.8	50.9	99014	6	ABN96931 Gene #342
c 36	152.8	50.9	110000	9	ACH03408_2 Continuation (3 of
c 37	152.8	50.9	112460	6	ABK83567 Human cDN
c 38	151.8	50.6	197997	10	AAAL54074 Human tra
c 39	151.6	50.5	246940	12	ADQ59422 Human can
c 40	151.4	50.5	76500	12	ADP45592 Human MAP
c 41	151.4	50.5	285020	11	ACN44958 Human gen
c 42	151.2	50.4	86131	10	ADF77178 KALPA Gen
c 43	151.2	50.4	114615	11	ACN44298 Human gen
c 44	151	50.3	499	4	AAI02649 Human rep
c 45	151	50.3	499	4	ABA07691 Human ova

ALIGNMENTS

RESULT 1  
ADH13943  
ID ADH13943 standard; DNA; 300 BP.

AC ADH13943;

DT 11-MAR-2004 (first entry)

DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:6.  
XX ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
KW cervical carcinoma.

OS Homo sapiens.

XX EP1170372-A1.

XX PD 09-JAN-2002.

XX PF 06-JUL-2000; 2000EP-00114560.

XX PR 06-JUL-2000; 2000EP-00114560.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Roehl F, Soto U, Coy J, Finzer P, Delius H, Poustka A;

XX Zur Hausen H, Patzelt A;

XX WPI; 2002-165895/22.

XX Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1, useful in gene therapy, for treating atherosclerosis and cancer.

XX Disclosure; SEQ ID NO 6; 30pp; English.

XX The invention relates to a novel nucleic acid molecule (I) comprising a sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein having the biological activity of MCP-1. A protein encoded by a nucleic acid of the invention has cytostatic, and antiarteriosclerotic activity. A nucleic acid of the invention may have a use in gene therapy. A compound of the invention is useful in the preparation of a medicament for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

CC pharmacological composition of the invention is useful for the treatment  
CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
CC atherosclerosis or cancer. The present sequence is used in the  
CC exemplification of the invention.

XX  
SQ Sequence 300 BP; 95 A; 60 C; 76 G; 69 T; 0 U; 0 Other;  
Query Match 100.0%; Score 300; DB 6; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.8e-77;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGCTAGTCCAGCTACTCGGGAGG 60  
Db 1 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGCTAGTCCAGCTACTCGGGAGG 60

QY 61 CTGAGGAGGAGAACTCTTTGAATCCAGGAGGCGCAGGTTGAGTGAGCAGAGATAGTGC 120  
Db 61 CTGAGGAGGAGAACTCTTTGAATCCAGGAGGCGCAGGTTGAGTGAGCAGAGATAGTGC 120

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180  
Db 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180

QY 181 AAAATGCGACGTGATTCAGCAGGCTCTGGGTTGAAGCCAGAACTCTCTGATAATTC 240  
Db 181 AAAATGCGACGTGATTCAGCAGGCTCTGGGTTGAAGCCAGAACTCTCTGATAATTC 240

QY 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTTAATAGAAAGCTACTGA 300  
Db 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTTAATAGAAAGCTACTGA 300

RESULT 2  
AD003803  
ID AD003803 standard; DNA; 11793 BP.  
XX  
AC AD003803;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human Cc12 gene and enhancer region DNA SeqID 4.  
XX  
KW human; ds; animal model; age-related macular degeneration; AMD;  
KW gene knockout; Cc12-deficient; Ccr2-deficient; drusen;  
KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;  
KW choroidal neovascularisation; ophthalmological; gene therapy.

XX  
OS Homo sapiens.  
XX  
PN WO2004041160-A2.  
XX  
PD 21-MAY-2004.  
XX  
PF 16-OCT-2003; 2003WO-US032933.  
XX  
PR 30-OCT-2002; 2002US-0422096P.  
XX  
PA (KENT ) UNIV KENTUCKY RES FOUND.  
XX  
PI Ambati J;  
XX  
DR WPI; 2004-400512/37.  
XX  
PT Testing candidate drug for treating age-related macular degeneration, by  
PT administering drug to Cc12-deficient, Ccr2-deficient knockout mouse, and  
PT analyzing development or regression of drusen and/or lipofuscin  
PT accumulation in eye.  
XX  
PS Disclosure; SEQ ID NO 4; 64pp; English.  
XX  
CC This invention relates to a novel methods and animal models for testing  
CC candidate drugs that can be used for the treatment or prevention of age-  
CC related macular degeneration (AMD). Specifically, it refers to

CC administering a candidate drug to gene knockout mice, in particular Cc12-  
CC deficient, Ccr2-deficient and/ or a Cc12-deficient/Ccr2-deficient dual  
CC knockout mouse. The present invention describes analysing the knockout  
CC mouse eye for development or regression of drusen and/ or lipofuscin  
CC accumulation, as well as for the effect of a candidate drug on Bruch's  
CC membrane, retinal degeneration and/ or choroidal neovascularisation.  
CC Accordingly, such compositions exhibit ophthalmological activities and  
CC can be used for gene therapy purposes. This polynucleotide sequence is  
CC the human Cc12 gene and enhancer region DNA of the invention.

XX  
SQ Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;  
Query Match 100.0%; Score 300; DB 12; Length 11793;  
Best Local Similarity 100.0%; Pred. No. 7.3e-77;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGCTAGTCCAGCTACTCGGGAGG 60  
Db 2500 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGCTAGTCCAGCTACTCGGGAGG 2559

QY 61 CTGAGGAGGAGAACTCTTTGAATCCAGGAGGCGCAGGTTGAGTGAGCAGAGATAGTGC 120  
Db 2560 CTGAGGAGGAGAACTCTTTGAATCCAGGAGGCGCAGGTTGAGTGAGCAGAGATAGTGC 2619

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180  
Db 2620 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 2679

QY 181 AAAATGCGACGTGATTCAGCAGGCTCTGGGTTGAAGCCAGAACTCTCTGATAATTC 240  
Db 2680 AAAATGCGACGTGATTCAGCAGGCTCTGGGTTGAAGCCAGAACTCTCTGATAATTC 2739

QY 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTTAATAGAAAGCTACTGA 300  
Db 2740 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTTAATAGAAAGCTACTGA 2799

RESULT 3  
ABL33652  
ID ABL33652 standard; DNA; 5926 BP.  
XX  
AC ABL33652;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human immune system associated gene SEQ ID NO: 1625.  
XX  
KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
KW ds.

XX  
OS Homo sapiens.  
XX  
PN WO200200928-A2.  
XX  
PD 03-JAN-2002.  
XX  
PF 02-JUL-2001; 2001WO-BF007537.  
XX  
PR 30-JUN-2000; 2000DE-01032529.  
XX  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
PA (EPIG-) EPIGENOMICS AG.  
XX  
PI Olek A, Piepenbrock C, Berlin K;  
XX  
DR WPI; 2002-130909/17.  
XX

PT Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.

XX Claim 1; SEQ ID NO 1625; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention

XX Sequence 5926 BP; 1562 A; 55 C; 1318 G; 2991 T; 0 U; 0 Other;

SQ Query Match 55.9%; Score 167.8; DB 6; Length 5926;

Best Local Similarity 80.7%; Pred. No. 2.2e-38;

Matches 196; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 58 AGCTCAGGCGAGAACCTCTGTAATCCAGGCGCGAGTTGCAGTCAGCAGATAG 117

Db 1 AGGTTGAGGTAGAGAAATTTTGAATTTAGGAGGCGTAGTGTGTAGTCAGTAGATAG 60

QY 118 TGCCACTGCACCTCCAGCTGGGTGCACAGTCAGACTCTCTCAAAAAATAAAATAAA 177

Db 61 TGTATTGTATTATTAGTTTGGGTAGATAGATGATTTGTTTAAAAAATAAATAAA 120

QY 178 ATAAAAATGCAGACTGTGATTCAGCAGCTCTGGGTGAAGCCAGAACTCTCTGATAAA 237

Db 121 ATAAAAATGTAGATTGTGATTTAGTAGGTTGGGTTGAAGTTTGAATTTTTCGATAAA 180

QY 238 TTCATGGCACTTAACCTACTTGGAGTTCATGGATGCTTTGCTAATCTAATAGAACCTAC 297

Db 181 TTTAATGGTATTAAATTATTGGAGGTTATGGATGTTTGTGTAATTTAATAGAAATTAT 240

QY 298 TGA 300

Db 241 TGA 243

RESULT 4

AAK67211/c

ID AAK67211 standard; DNA; 12026 BP.

XX AC AAK67211;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22023.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR PT-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.

CC supplement the patients own production of (I). Additionally, (I) CC  
CC polynucleotides may be used to produce the secreted (I), by inserting the CC  
CC nucleic acids into a host cell and culturing the cell to express the CC  
CC protein. (I) proteins and polynucleotides may be used to prevent, CC  
CC diagnose and treat immune/haematopoietic-related diseases, especially CC  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703 CC  
CC to AAK87694 represent human immune/haematopoietic antigen genomic CC  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169 CC  
CC represent sequences used in the exemplification of the present invention CC  
XX  
SQ Sequence 12026 BP; 2791 A; 3406 C; 3411 G; 2418 T; 0 U; 0 Other;  
Query Match 53.6%; Score 160.8; DB 4; Length 12026;  
Best Local Similarity 86.8%; Pred. No. 31e-36;  
Matches 177; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
QY 1 AAAATATAAAAATTAGCCAGGCGTGATGTCATGCTGCTGTAGTCCAGCTACTCCGGAGG 60  
Db 4683 AAAATACAAAATTAGCCAGGCGTGATGTCATGCTGCTGTAGTCCAGCTACTCCGGAGG 4624  
QY 61 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGAGGTTGTCAGTGAGCAGAGATAGTGC 120  
Db 4623 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGAGGTTGTCAGTGAGCAGAGATAGTGC 4564  
QY 121 CACTGCACCTCCAGCCTCGGTGACAGAGTGAGACTCTGTCTCAAAAATAAATAAATA 180  
Db 4563 CACTGCACCTCCAGCCTCGGTGACAGAGTGAGACTCTGTCTCAAAAATAAATAAATA 4504  
QY 181 AAAATGCAGACTGTGATTCAGCA 204  
Db 4503 GAAATGGCTAAGTGTGAGGCAGAA 4480  
RESULT 5  
ABD33528  
ID ABD33528 standard; DNA; 31926 BP.  
XX  
AC ABD33528;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human cancer-associated (CA) gene HD07-104.  
XX  
KW Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;  
XX ds; cancer; cytostatic.  
XX  
OS Homo sapiens.  
XX  
FN WO2004058146-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 15-DEC-2003; 2003WO-US040081.  
XX  
PR 17-DEC-2002; 2002US-00322281.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Malandro MS;  
XX  
WPI; 2004-499109/47.  
XX  
PT Novel human cancer associated protein encoded within open reading frame  
XX of cancer associated gene, useful as targets for diagnosing cancer.  
XX  
PS Claim 16; SEQ ID NO 712; 182pp; English.  
XX  
CC The invention relates to cancer-associated proteins (CAP) and the cancer-  
CC associated (CA) nucleic acids encoding them. The invention also relates  
CC to a method for treating cancers involving administering to a patient an  
CC inhibitor of CAP, and a method of screening for anticancer activity in a  
CC potential drug involving providing a cell that expresses a CA gene,  
CC contacting a tissue sample derived from a cancer cell with an anticancer  
CC

PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
PT  
XX  
XX Disclosure; SEQ ID NO 22023; 3071pp + Sequence Listing; English.  
PS  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to

CC drug candidate and monitoring the effect of the anticancer drug candidate  
CC on expression of the CA gene. The CAP proteins are useful for detecting  
CC cancer associated with expression of a CAP protein in a test cell sample  
CC and for screening for a bioactive agent capable of modulating the  
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
CC cancer, involving determining the expression of a CA nucleic acid in a  
CC tissue. This sequence represents a human CA gene of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 31926 BP; 6797 A; 9017 C; 8858 G; 7254 T; 0 U; 0 Other;

Query Match 53.6%; Score 160.8; DB 13; Length 31926;

Best Local Similarity 86.8%; Pred. No. 4.5e-36;

Matches 177; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 AAAATAAAATACCGAGCGGTGATGTCATGTGCTGTAGTCCAGTACTTCGGAGG 60  
DB 17249 AAAATAAAATACCGAGCGGTGATGTCATGTGCTGTAGTCCAGTACTTCGGAGG 17308

QY 61 CTGAGCAGGAGAACCTCTTGATCCAGGCGCAGGTTCAGTGCAGAGATAGTGC 120  
DB 17309 CTGAGCAGGAGAACCTCTTGATCCAGGCGCAGGTTCAGTGCAGAGATAGTGC 17368

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180  
DB 17369 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 17428

QY 181 AAAATGCACTGTGATTCAGCA 204  
DB 17429 GAAATGCTAAGTGTGAGGCAGAA 17452

RESULT 6

ID ABK42328/c

XX ABK42328 standard; DNA; 5304 BP.

AC ABK42328;

XX DT 21-MAY-2002 (first entry)

XX Genomic sequence #227 encoding novel human connective tissue polypeptide.

XX Human; connective tissue related disorder; cancer; gene therapy;

KW cytosstatic; gene; ds.

XX OS Homo sapiens.

XX PN WO200155343-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-05001322.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225477P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

Best Local Similarity 81.5%; Pred. No. 4.4e-35; Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60  
Db 3630 AAACATATAAAATAGCTGGCGTGGTGGCACGTGCTGTAGTCCAGCTACTCGGGAGG 3571  
QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGGCGAGGTTGCAGTCAGCAGAGATAGTGC 120  
Db 3570 CTGAGGCAGAGAAATCGCTTTAAACCCGGGAGGGGAGATTGCAGTGAGCTGAGATCATGC 3511  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATA 180  
Db 3510 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATA 3451  
QY 181 AAAAATGCAGACTGTGATTCAGCAGGCTCTGGTGTGAAGGCCA 222  
Db 3450 AAGAAGAAGAGAAAGAGAGGCTGTCTCTCGGTCAAGGCCA 3409

RESULT 7  
ABK42327/c  
ID ABK42327 standard; DNA; 5304 BP.  
AC ABK42327;  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE Genomic sequence #226 encoding novel human connective tissue polypeptide.  
XX  
KW Human; connective tissue related disorder; cancer; gene therapy;  
KW cytosstatic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155343-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001322.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205513P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.

08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-02559678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-565190/63.  
DR  
XX Nucleic acid encoding novel connective tissue associated polypeptides,  
PT used in diagnosing, preventing, treating or ameliorating a disorder such  
PT as cancer or rheumatoid arthritis.  
XX  
PS Disclosure; SEQ ID NO 1215; 673pp; English.  
XX  
XX The present invention relates to the isolation of novel human connective  
CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide  
CC (cDNA and genomic) sequences encoding them. The sequences of the  
CC invention are useful in the diagnosis, treatment, prevention and/or  
CC prognosis of diseases associated with connective tissue(s), including  
CC cancer. The polynucleotide sequences of the invention are also useful in  
CC gene therapy. ABK42102-ABK43116 represent genomic sequences encoding the  
CC novel human connective tissue related polypeptides. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 5304 BP; 1550 A; 1037 C; 1017 G; 1700 T; 0 U; 0 Other;  
Query Match 52.1%; Score 156.4; DB 4; Length 5304;

```
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231141P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232401P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-565190/63.
XX
XX Nucleic acid encoding novel connective tissue associated polypeptides,
XX used in diagnosing, preventing, treating or ameliorating a disorder such
XX as cancer or rheumatoid arthritis.
XX
XX Disclosure; SEQ ID NO 1214; 673pp; English.
XX
XX The present invention relates to the isolation of novel human connective
XX tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
XX (cDNA and genomic) sequences encoding them. The sequences of the
XX invention are useful in the diagnosis, treatment, prevention and/or
XX prognosis of diseases associated with connective tissue(s), including
XX cancer. The polynucleotide sequences of the invention are also useful in
XX gene therapy. ABK42102-ABK43116 represent genomic sequences encoding the
XX novel human connective tissue related polypeptides. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 5304 BP; 1550 A; 1038 C; 1017 G; 1699 T; 0 U; 0 Other;
XX
XX Query Match 52.1%; Score 156.4; DB 4; Length 5304;
XX Best Local Similarity 81.5%; Pred. No. 4.4e-35;
XX Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
XX
XX 1 AAATATATAAAATTAGCCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCCGGAGG 60
XX 3630 AAACCTATATAAAATTAGCTGGCGGTGGTGGACGCTGCTAGTCCAGCTACTCCGGAGG 3571
XX
XX 61 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGCGCAGGTTCGAGTGAGCAGAGATAGTGC 120
XX 3570 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGCGCAGGTTCGAGTGAGCTGAGATCATGC 3511
XX
XX 121 CACTGCACCTCCAGCCTGGGTGACAGGTGAGACTCTGTCTCAAAAAATAAAATAATA 180
XX 3510 CACTGCACCTCCAGCCTGGGTGACAGGTGAGACTCTGTCTCAAAAAATAAAATAATA 3451
```

[illegible]

PR	17-NOV-2000;	2000US-0249207P.	
PR	17-NOV-2000;	2000US-0249208P.	
PR	17-NOV-2000;	2000US-0249209P.	
PR	17-NOV-2000;	2000US-0249210P.	
PR	17-NOV-2000;	2000US-0249211P.	
PR	17-NOV-2000;	2000US-0249212P.	
PR	17-NOV-2000;	2000US-0249213P.	
PR	17-NOV-2000;	2000US-0249214P.	
PR	17-NOV-2000;	2000US-0249215P.	
PR	17-NOV-2000;	2000US-0249216P.	
PR	17-NOV-2000;	2000US-0249217P.	
PR	17-NOV-2000;	2000US-0249218P.	
PR	17-NOV-2000;	2000US-0249244P.	
PR	17-NOV-2000;	2000US-0249245P.	
PR	17-NOV-2000;	2000US-0249264P.	
PR	17-NOV-2000;	2000US-0249265P.	
PR	17-NOV-2000;	2000US-0249267P.	
PR	17-NOV-2000;	2000US-0249297P.	
PR	17-NOV-2000;	2000US-0249299P.	
PR	17-NOV-2000;	2000US-0249300P.	
PR	01-DEC-2000;	2000US-0250160P.	
PR	01-DEC-2000;	2000US-0250391P.	
PR	05-DEC-2000;	2000US-0251030P.	
PR	05-DEC-2000;	2000US-0251988P.	
PR	05-DEC-2000;	2000US-0256719P.	
PR	06-DEC-2000;	2000US-0251479P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
PR	08-DEC-2000;	2000US-0251989P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
PR	17-JAN-2001;	2001US-00764847.	
XX	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
XX	Rosen CA, Ruben SM, Barash SC;		
XX			
XX	WPI; 2003-634869/60.		
XX	P-PSDB; ADB59409.		
XX			
PT	New connective tissue-related polypeptides and polynucleotides, useful		
PT	for treating, preventing and/or prognosing e.g. disorders of connective		
PT	tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or		
PT	neoplasias.		
XX			
PS	Disclosure; SEQ ID NO 1214; 248pp; English.		
XX			
CC	The invention describes an isolated nucleic acid molecule (I), which		
CC	comprises a sequence that is at least 95 % identical to a connective		
CC	tissue-related polynucleotide encoding connective tissue antigens (CTA).		
CC	The polypeptide or polynucleotide is useful for preventing, treating, or		
CC	ameliorating medical conditions in a mammal. The connective tissue		
CC	polypeptides, polynucleotides and antibodies are particularly useful for		
CC	treating, preventing and/or prognosing disorders of connective tissues		
CC	(e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,		
CC	scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or		
CC	neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.		
CC	Alzheimer's disease, or Parkinson's disease), cardiovascular diseases		
CC	(e.g. atherosclerosis, myocarditis or cardiopulmonary bypass		
CC	complications), autoimmune diseases (e.g. systemic lupus erythematosus,		
CC	rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.		
Query Match 52.1%; Score 156.4; DB 9; Length 5304;			
Best Local Similarity 81.5%; Pred. No. 4.4e-35;			
Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;			
QY	1	AAAAATAAAATTAGCCAGCGCTGATGTGTCATGCTGTAGTCCAGCTACTTCGGGAGG 60	
Db	3630	AAACTATAAAATTAGTGGCGTGTGGCACGTGCTGTAGTCCAGCTACTTCGGGAGG 3571	
QY	61	CTGAGGAGGAGAACTTTGAATCCAGGAGCGCGAGTTGTCAGTGACGAGATAGTGC 120	

Db	3570	CTGAGGAGGAGAAATCGCTTAAACCCGGGAGCGGAGATTGCAGTGTGATCATGC 3511	
QY	121	CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATA 180	
Db	3510	CACTGCACCTCCAGCTGGGCGACAGAGTGAGGCTCTGCTCTCAAAAAATAAAAAATA 3451	
QY	181	AAAAATGCAGACTGTGATTTCAGCAGGTTCTGGGTTGAAGCCCA 222	
Db	3450	AAAGAGAGAGAGAGAGAGAGGCTGTCTCTGTCGTCAGGCCCA 3409	
RESULT 9			
ADB60484/c			
ID	ADB60484	standard; DNA; 5304 BP.	
XX			
AC	ADB60484;		
XX			
DT	04-DEC-2003	(first entry)	
XX			
DE		Connective tissue related genomic DNA #227.	
XX			
KW		cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;	
KW		antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;	
KW		antiinflammatory; antiallergic; antiasthmatic; dermatological;	
KW		nephrotropic; virucide; fungicide; antibacterial; antiparasitic;	
KW		gene therapy; ds; connective tissues disorder; rheumatoid arthritis;	
KW		systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;	
KW		cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;	
KW		Alzheimer's disease; Parkinson's disease; cardiovascular disease;	
KW		atherosclerosis; myocarditis; cardiopulmonary bypass complication;	
KW		autoimmune disease; multiple sclerosis; allergic reaction; asthma;	
KW		rinitis; eczema; inflammatory condition; Crohn's disease; nephritis;	
KW		gastrointestinal disorder; inflammatory bowel disease;	
KW		organ transplant rejection; immune system disorder; Bruton's disease;	
KW		X-linked lymphoproliferative syndrome;	
KW		B-cell lymphoproliferative disorder; HIV; AIDS; infection;	
KW		chromosome identification; chromosome mapping;	
KW		connective tissue related polynucleotide; gene; ds.	
XX			
OS		Homo sapiens.	
XX			
PN		US2003054375-A1.	
XX			
PD		20-MAR-2003.	
XX			
PF		07-MAR-2002; 2002US-00092154.	
XX			
PR		31-JAN-2000; 2000US-0179065P.	
PR		04-FEB-2000; 2000US-0180628P.	
PR		24-FEB-2000; 2000US-0184664P.	
PR		02-MAR-2000; 2000US-0186350P.	
PR		16-MAR-2000; 2000US-0189874P.	
PR		17-MAR-2000; 2000US-0190076P.	
PR		18-APR-2000; 2000US-0198123P.	
PR		19-MAY-2000; 2000US-0205515P.	
PR		07-JUN-2000; 2000US-0209467P.	
PR		28-JUN-2000; 2000US-0214886P.	
PR		30-JUN-2000; 2000US-0215135P.	
PR		07-JUL-2000; 2000US-0216647P.	
PR		07-JUL-2000; 2000US-0216880P.	
PR		11-JUL-2000; 2000US-0217487P.	
PR		14-JUL-2000; 2000US-0218290P.	
PR		26-JUL-2000; 2000US-0220963P.	
PR		26-JUL-2000; 2000US-0220964P.	
PR		14-AUG-2000; 2000US-0224518P.	
PR		14-AUG-2000; 2000US-0224519P.	
PR		14-AUG-2000; 2000US-0225213P.	
PR		14-AUG-2000; 2000US-0225214P.	
PR		14-AUG-2000; 2000US-0225266P.	
PR		14-AUG-2000; 2000US-0225267P.	
PR		14-AUG-2000; 2000US-0225268P.	
PR		14-AUG-2000; 2000US-0225270P.	

PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244675P.  
PR 08-NOV-2000; 2000US-0244676P.  
PR 08-NOV-2000; 2000US-0244677P.  
PR 08-NOV-2000; 2000US-0244678P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764847.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Barash SC;

WPI; 2003-634869/60.

P-PSDB; ADB59409.

PT New connective tissue-related polypeptides and polynucleotides, useful for treating, preventing and/or prognosing e.g. disorders of connective tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or neoplasias.

XX Disclosure; SEQ ID NO 1215; 248pp; English.

XX The invention describes an isolated nucleic acid molecule (I), which comprises a sequence that is at least 95 % identical to a connective tissue-related polynucleotide encoding connective tissue antigens (CTA). The polypeptide or polynucleotide is useful for preventing, treating, or ameliorating medical conditions in a mammal. The connective tissue polypeptides, polynucleotides and antibodies are particularly useful for treating, preventing and/or prognosing disorders of connective tissues (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus, scleroderma or Sjogren's syndrome), cancers, cancer metastases and/or neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g. Alzheimer's disease, or Parkinson's disease), cardiovascular diseases (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass complications), autoimmune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.

Query Match 52.1%; Score 156.4; DB 9; Length 5304;  
Best Local Similarity 81.5%; Pred. No. 4.4e-35;  
Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGCGTGATGTCATGCTGTAGTCCCGAGTACTCGGAGG 60

Db 3630 AAACATAAAATAGTGGCGGTGGTGGCACGTGCTGTAGTCCAGCTACTCGGGAGG 3571  
QY 61 CTGAGCGAGAGAACCTCTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGATAGTGC 120  
Db 3570 CTGAGCGAGAGAAATCGCTTAAACCCGGGAGCGGAGATTGCAGTGCAGCTGAGATCATGC 3511  
QY 121 CACTGCACCTCCAGCCTGGGTGCAGAGTGCAGAGTCTGTCTCAAAAAATAAAATAAATA 180  
Db 3510 CACTGCACCTCCAGCCTGGGTGCAGAGTGCAGAGTCTGTCTCAAAAAATAAAATAAATA 3451  
QY 181 AAAAATGCAGACTGTGATTCAGCAGAGTCTGGGTGAAGGCCA 222  
Db 3450 AAAGAAGAAGAAGAAGAAGGCTGTCTGTGTCAGGCCA 3409

## RESULT 10

ABK69846  
ID ABK69846 standard; DNA; 9796 BP.

XX ABK69846;

DT 15-JUL-2002 (first entry)

XX Human secreted protein gene 22 genomic DNA fragment #7.

XX Human; ds; secreted protein; gene therapy; immunosuppressive;  
XX antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
KW virucide; fungicide; opthalmological; autoimmune disease; neoplasm;  
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; wound healing; skin aging;  
KW epithelial cell proliferation; food additive.

XX Homo sapiens.

XX WO200226931-A2.

XX 04-APR-2002.

XX 24-SEP-2001; 2001WO-US029871.

XX 25-SEP-2000; 2000US-0234925P.

PR 12-JAN-2001; 2001WO-US000911.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;  
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;  
PI Endress GA, Mucenski M, Ebner R;

XX WPI; 2002-362489/39.

XX Novel 71 isolated secreted polypeptides and polynucleotides encoding the  
PT polypeptides, useful for treating Huntington's disease, sepsis,  
PT meningitis, thrombocytopenia, hemolytic anemia, rheumatoid arthritis,  
PT asthma.

XX Example 2; Page 1381-1383; 1478pp; English.

XX The invention relates to an isolated nucleic acid molecule (or its  
CC fragment, homologue complement or allelic variant) encoding a human  
CC secreted protein (and its fragment, domain, epitope, variant, secreted  
CC form and species variant). Also included are a recombinant vector  
CC comprising the nucleic acid, a recombinant host cell comprising the  
CC vector, an antibody against the secreted protein, a recombinant host cell  
CC that expresses the secreted protein and a method of identifying a binding  
CC partner of the secreted protein. The nucleic acid and protein are used to  
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.  
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep for  
CC example autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. Many other diseases and disorders are listed in  
CC the specification. The polypeptides can also be used to aid wound healing  
CC an epithelial cell proliferation, to prevent skin aging due to sunburn,  
CC to maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. The present sequence  
CC represents a ds DNA fragment of the gene for a novel human secreted  
CC protein of the invention  
XX

XX Sequence 9796 BP; 2064 A; 2679 C; 2566 G; 2487 T; 0 U; 0 Other;

Query Match 52.1%; Score 156.4; DB 6; Length 9796;

Best Local Similarity 85.0%; Pred. No. 5.5e-35;

Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60

Db 5808 AAAATACAGAAATTAGTGGGTGTGATGTCGTCCTGTATCCAGCTACTCGGGAGC 5867

QY 61 CTGAGCGAGGAGAACCTCTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGATAGTGC 120

Db 5868 CTGAGCGAGGAGAACCTCTTGAATCCGGAGCGGAGGTTGCAGTGCAGCGAGATTGTAC 5927

QY 121 CACTGCACCTCCAGCCTGGGTGCAGAGTGCAGTGCAGTCTGTCTCAAAAAATAAAATA 180

Db 5928 CACTGCACCTCCAGCCTGGGCCACAGAGAAAGACTGTCTCTCAAAAAATAAAATA 5987

QY 181 AAAAATGCAGACTGTGATTCAGCAGG 206

Db 5988 AAAAATCCAGCCTGAGTCTCACCAGG 6013

## RESULT 11

ACC50861

ID ACC50861 standard; cDNA; 9796 BP.

XX ACC50861;

XX 12-JUN-2003 (first entry)

XX Human secreted protein BAC clone SEQ ID NO 1041.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;  
KW vulnery; antiinflammatory; neurotropic; neuroprotective;  
KW antiparkinsonian; gene therapy; human; cardiovascular disorder; ss.

OS Homo sapiens.

XX WO200295010-A2.

XX 28-NOV-2002.

XX 19-MAR-2002; 2002WO-US0009785.

XX 21-MAR-2001; 2001US-0277340P.

PR 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-129429/12.

XX Novel human secreted proteins, useful for detecting, preventing,  
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular  
PT disorders such as arrhythmia.

Disclosure; SEQ ID NO 1041; 1881pp; English.

PS The present invention relates to novel human secreted proteins (ABR47633-  
XX ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins  
CC and their coding sequences are useful for the preparation of a diagnostic  
CC or pharmaceutical composition for diagnosing or treating a cardiovascular  
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary  
CC arteriosclerosis and myocardial ischaemia), neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
CC proliferative disorders and/or cancerous diseases and conditions, for  
CC wound healing and epithelial cell proliferation, to treat inflammation or  
CC infection, for treating thrombosis and arteriosclerosis, for treating or  
CC preventing neural damage which occurs in neuronal disorders or  
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
CC disease, to enhance bone and periodontal regeneration and aid in tissue  
CC transplants or bone grafts, to prevent skin aging or hair loss, to  
CC stimulate growth and differentiation of haematopoietic cells and bone  
CC marrow cells when used in combination with other cytokines, to maintain  
CC organs before transplantation or for supporting cell culture of primary  
CC tissues, to increase or decrease differentiation or proliferation of  
CC embryonic stem cells, or to modulate mammalian characteristics or  
CC metabolism. The present sequence was used to illustrate the invention.  
CC Note: The sequence data for this patent was published in electronic  
CC format and is available from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 9796 BP; 2064 A; 2679 C; 2566 G; 2487 T; 0 U; 0 Other;

Query Match 52.1%; Score 156.4; DB 8; Length 9796;  
Best Local Similarity 85.0%; Pred. No. 5.5e-35;  
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
QY 1 AAAATATAAAATTTAGCCAGCGGTGATGTCATGTGCTGTGATGCCAGCTACTCCGGAGG 60  
DB 5808 AAAATACAGAAATTAGCTGGGTGTGATGGTGGCTGTGTAATCCAGCTACTCCGGAGG 5867  
QY 61 CTGAGGCGAGGAGAACCTCTTTGAATCCAGGAGGCGGAGGTTGCAGTGACGAGATAGTGC 120  
DB 5868 CTGAGGCGAGGAGAACCTCTTTGAATCCAGGAGGCGGAGGTTGCAGTGACGAGATAGTGC 5927  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATAAATA 180  
DB 5928 CACTGCACCTCCAGCTGGGCGCACAGAAAGACTCTGTCTCAAAAAATAAAAAATAAATA 5987  
QY 181 AAAAATGCAGACTGTGATTCAGCAGG 206  
DB 5988 AAAAATCCAGCTGAGTCTCACCAGG 6013

RESULT 12  
ABZ71483  
ID ABZ71483 standard; DNA; 9796 BP.

XX AC ABZ71483;

XX DT 03-APR-2003 (first entry)

XX DE Secreted protein gene 12 genomic fragment HAUAI83, SEQ ID NO:593.

XX KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
KW immune disorder; inflammation; infection; wound healing; drug screening;  
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
KW antiinflammatory; immunosuppressive; vulnerability; chromosome 19; gene; ds.

XX OS Homo sapiens.

XX PN WO200276488-A1.

XX PD 03-OCT-2002.

PF 19-MAR-2002; 2002WO-US008276.

XX 21-MAR-2001; 2001US-0277340P.

PR 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-029900/02.

XX New human secreted proteins and nucleic acids, useful for detecting,  
XX preventing, diagnosing, prognosticating, treating and/or ameliorating  
XX e.g. gastrointestinal diseases and disorders, or cancers.

PS Disclosure; Page 1143-1145; 1216pp; English.

XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted  
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The  
CC invention also encompasses antibodies specific for the secreted proteins,  
CC the use of the secreted proteins in drug screening, and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC fragments specific for the secreted proteins, and modulators of protein  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
CC activity are useful for diagnosing, treating, ameliorating or preventing  
CC digestive disorders. Such conditions include disorders of the mouth,  
CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
CC tract and pancreas, and include cancers of these organs and tissues. The  
CC secreted proteins and their nucleic acids may also be used in the  
CC treatment of immune disorders, inflammation, infection,  
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
CC of the invention may be used for chromosome identification, chromosome  
CC mapping, in gene therapy, for identifying individuals from minute  
CC biological samples, as hybridisation probes, and as molecular weight  
CC markers. The present sequence represents a human secreted protein genomic  
CC fragment referred to in the disclosure of the invention

XX SQ Sequence 9796 BP; 2064 A; 2679 C; 2566 G; 2487 T; 0 U; 0 Other;

Query Match 52.1%; Score 156.4; DB 8; Length 9796;  
Best Local Similarity 85.0%; Pred. No. 5.5e-35;  
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTTAGCCAGCGGTGATGTCATGTGCTGTGATGCCAGCTACTCCGGAGG 60  
DB 5808 AAAATACAGAAATTAGCTGGGTGTGATGGTGGCTGTGTAATCCAGCTACTCCGGAGG 5867

QY 61 CTGAGGCGAGGAGAACCTCTTTGAATCCAGGAGGCGGAGGTTGCAGTGACGAGATAGTGC 120  
DB 5868 CTGAGGCGAGGAGAACCTCTTTGAATCCAGGAGGCGGAGGTTGCAGTGACGAGATAGTGC 5927

QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAAAATAAATA 180  
DB 5928 CACTGCACCTCCAGCTGGGCGCACAGAAAGACTCTGTCTCAAAAAATAAAAAATAAATA 5987

QY 181 AAAAATGCAGACTGTGATTCAGCAGG 206

DB 5988 AAAAATCCAGCTGAGTCTCACCAGG 6013

RESULT 13

ADB91839

ID ADB91839 standard; DNA; 9796 BP.

XX AC ADB91839;

XX DT 04-DEC-2003 (first entry)

XX DE Human secreted protein related DNA #SEQ ID 785.

XX KW Secreted protein; gene therapy; antidiabetic; diabetes; human; ds.

```
XX OS Homo sapiens.
XX KW W02003004622-A2.
XX PN
XX KW
XX PD 16-JAN-2003.
XX PF 19-MAR-2002; 2002WO-US008124.
XX PR 21-MAR-2001; 2001IUS-0277340P.
XX PR 19-JUL-2001; 2001IUS-0306171P.
XX PR 13-NOV-2001; 2001IUS-0331287P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX XX WPI; 2003-229407/22.
XX DR
XX PT Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX PT treating diabetes or conditions related to diabetes.
XX PS Disclosure; SEQ ID NO 785; 1537pp; English.
XX CC The invention relates to isolated nucleic acid molecules ADB91065-
XX CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX CC ADB91834. Also disclosed is a recombinant vector comprising a
XX CC polynucleotide of the invention, and a recombinant host cell comprising
XX CC the recombinant vector. The polypeptide of the invention is useful in
XX CC identifying a binding partner by contacting the polypeptide with a
XX CC binding partner, and determining whether the binding partner increases or
XX CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX CC antibody or its fragment, agonist or antagonist are useful for preparing
XX CC a pharmaceutical composition for diagnosing or treating diabetes or
XX CC conditions related to diabetes. The present sequence is that of the human
XX CC immunoglobulin Fc portion used to generate fusion proteins, increasing
XX CC the stability of the fused protein as compared to the secreted protein
XX CC only. Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 9796 BP; 2064 A; 2679 C; 2566 G; 2487 T; 0 U; 0 Other;

Query Match 52.1%; Score 156.4; DB 9; Length 9796;
Best Local Similarity 85.0%; Pred. No. 5.5e-35;
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATATAAATAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 60
DB 5808 AAAATACAGAAATAGCTGGGTGATGTCGTCGTGCTGTAATCCAGCTACTCGGGAGC 5867
QY 61 CTGAGGCGAGGAGAACCTTTGAATCCAGGCGGAGGTTGAGTGCAGCAGAGATAGTGC 120
DB 5868 CTGAGGCGAGGAGAACCTTTGAATCCAGGCGGAGGTTGAGTGCAGGAGATTGTAC 5927
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGCAGTCTGTCTCAAAAAATAAAATAAATA 180
DB 5928 CACTGCACCTCCAGCTGGGCCACAGAGAAAGACTCTGTCTCAAAAAATAAAATAAATA 5987
QY 181 AAAATGCAGACTGTGATTACAGCAGG 206
DB 5988 AAAATCCAGCTGAGTCTCACCAGG 6013

RESULT 14
ADC74614
ID ADC74614 standard; DNA; 9796 BP.
XX AC ADC74614;
XX XT
XX DT 01-JAN-2004 (first entry)
XX DE Human secreted protein-related DNA - SEQ ID 1247.
```

```
XX KW antianaemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
XX KW antidiabetic; immunosuppressive; dermatological; nephroretropic;
XX KW antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
XX KW fungicide; antiparasitic; antiarteriosclerotic; vulnerrary; cytostatic;
XX KW haemopoietic; haematologic; anaemia; autoimmune disorder;
XX KW rheumatoid arthritis; inflammation; Grave's disease; diabetes;
XX KW systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
XX KW Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
XX KW cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
XX KW human; ds.
XX OS Homo sapiens.
XX PN W02003038063-A2.
XX XX
XX PD 08-MAY-2003.
XX PF 19-MAR-2002; 2002WO-US008277.
XX PR 21-MAR-2001; 2001IUS-0277340P.
XX PR 19-JUL-2001; 2001IUS-0306171P.
XX PR 13-NOV-2001; 2001IUS-0331287P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX XX WPI; 2003-430516/40.
XX DR
XX PT New human secreted polypeptide for diagnosing, preventing or treating
XX PT hemopoietic or hematologic disorders (e.g. anemia), autoimmune
XX PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
XX PT atherosclerosis).
XX PS Disclosure; SEQ ID NO 1247; 2272pp; English.
XX CC The invention relates to a novel human secreted polypeptide comprising a
XX CC defined sequence given in the specification. The polypeptide, nucleic
XX CC acid molecule, antibody, agonist or antagonist of the invention may be
XX CC useful for preparing a composition for diagnosing or treating a
XX CC haemopoietic or haematologic disorder such as anaemia, autoimmune
XX CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
XX CC diabetes, systemic lupus erythematosus or glomerulonephritis,
XX CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
XX CC disease, wounds and hyperproliferative disorders including
XX CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
XX CC parasitic infections. The polypeptide may also be used during gene
XX CC therapy procedures and for identifying a binding partner by contacting
XX CC the polypeptide with a binding partner and determining whether the
XX CC binding partner increases or decreases the activity of the polypeptide.
XX CC The current sequence is that of the human secreted protein-related DNA of
XX CC the invention.
XX SQ Sequence 9796 BP; 2064 A; 2679 C; 2566 G; 2487 T; 0 U; 0 Other;

Query Match 52.1%; Score 156.4; DB 10; Length 9796;
Best Local Similarity 85.0%; Pred. No. 5.5e-35;
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATATAAATAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 60
DB 5808 AAAATACAGAAATAGCTGGGTGATGTCGTCGTGCTGTAATCCAGCTACTCGGGAGC 5867
QY 61 CTGAGGCGAGGAGAACCTTTGAATCCAGGCGGAGGTTGAGTGCAGCAGAGATAGTGC 120
DB 5868 CTGAGGCGAGGAGAACCTTTGAATCCAGGCGGAGGTTGAGTGCAGGAGATTGTAC 5927
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGCAGTCTGTCTCAAAAAATAAAATAAATA 180
DB 5928 CACTGCACCTCCAGCTGGGCCACAGAGAAAGACTCTGTCTCAAAAAATAAAATAAATA 5987
QY 181 AAAATGCAGACTGTGATTACAGCAGG 206
DB 5988 AAAATCCAGCTGAGTCTCACCAGG 206
```

Qy	61	CTGAGGCAGGAGAACTCTTTGAATCCAGGAGGCGCAGGTTGCAGTGACGACGAGATAGTGC	120
Db	5868	CTGAGGCAGGGGAATCTCTTTGAATCCGGGAGCGGAGGTTGCAGTGACGCGAGATTCTAC	5927
Qy	121	CACGTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTCTCTCAAAAAAATATAAATAAATA	180
Db	5928	CACGTGCACCTCCAGCCTGGGCCACAGAGAAAGACTCTGTCTCAAAAAAATAAAAAAAA	5987
Qy	181	AAAAATGCAGACTGTGATTCAGCAGG	206
Db	5988	AAAAATCCAGCCTCAGTCTCACCAGG	6013

Search completed: August 4, 2005, 08:31:24  
Job time : 203.826 secs

5988 AAAATCCAGCCTGAGTCTCACCAGG 6013

RESULT 15  
ADD38134  
IN ADD38134 standard: CDNA: 9796 BP.

AC	ADD38134;
XX	
DT	15-JAN-2004 (first entry)

cdNA clone in ATCC deposit #28.

human secreted protein; Antiallergic; Antiinflammatory; Antibacterial; Antihiv. Catestatic. Immunosuppressive; Hemostatic; ss.

**Homo sapiens.**

WO200290526-A2.

14-NOV-2002.

19-MAR-2002:

21-MAR-2001: 2001US-0277340P.

19-00F-2001; 2001US-0331287P;  
13-NOV-2001; 2001US-0331287P.

(HITMA-) HITMAN GENOME SCI INC.

BOGGS CA      PUGHEN SM.

WPT: 2003-140218/13.

New human secreted m

preparing a diagnosis or treating allergic or

hyper-sensitivity disorders.

Claim 1. SEQ ID NO 616: 1323pp: English.

The present invention relates to an isolated polypeptide or human secreted protein. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders. The polypeptide is also useful for identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases the activity of the polypeptide. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating inflammatory disorders, neoplastic diseases, wound healing and disorders of epithelial cell proliferation, immune disorders, cardiovascular disorders, blood-related disorders, infectious diseases, endocrine disorders, or gastrointestinal disorders. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The present sequence represents a cDNA clone from ATCC deposit.

XX  
CO  
commence 9796 BP. 2064 A: 2679 C: 2566 G: 2487 T: 0 U; 0 Other;

```

Query Match      52.1%; Score 156.4; DB 10; Length 9796;
Best Local Similarity 85.0%; Pred. No. 5.5e-35;
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```

Qy 1 A A A A T A T A A A A T T A G C C A G G C G T G A T C A T G T G C C T G T A G T C C C A G T A C T C G G A G G 60

Dh 5808 A A A A T A C A G A A A T T A G C T G G G T G T G A T G G T G C G T G C C T A T C C A G C T A T C G G G A G C 5867

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 4, 2005, 04:04:35 ; Search time 1415.95 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-6  
Perfect score: 300  
Sequence: 1 aaataataaaattagccag.....aatctaataagaagctactga 300

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	154.8	51.6	828	8	AQ749749	HS 5573 A
C 2	151	50.3	344	2	BF871137	CM0-ET012
C 3	151	50.3	526	7	CR556387	DKFZP459P
C 4	150.6	50.2	800	8	AQ750780	HS_5576_B
C 5	150.4	50.1	675	8	AQ313572	RPCI11-10
C 6	150.4	50.1	757	8	BZ602060	WHABV68TR
C 7	150.2	50.1	565	8	AQ474644	CITBI-EI-
C 8	150.2	50.1	657	8	AQ490406	RPCI-11-2
C 9	150	50.0	751	7	CF128305	UI-HF-ET0
C 10	149.8	49.9	491	2	AW410354	fh05f06.y
C 11	149.6	49.9	572	8	AQ268868	RPCI11-73
C 12	149.6	49.9	638	8	AQ312235	RPCI11-10
C 13	149.4	49.8	642	9	AG047289	Pan trogl
C 14	149.4	49.8	677	9	AG160213	Pan trogl
C 15	149.4	49.8	732	8	AQ890095	HS 3100_A
C 16	149.2	49.7	526	8	AQ358759	HS_5030_A
C 17	149	49.7	619	5	BX470875	DKFZP686D
C 18	148.8	49.6	435	1	A1460009	ac93f05.x
C 19	148.8	49.6	497	7	CF140288	UI-HF-CB0
C 20	148.8	49.6	629	5	BX507857	DKFZP686D
C 21	148.8	49.6	766	5	BQ000069	UI-H-DP0-
C 22	148.8	49.6	771	5	BQ776208	UI-H-FH0-
C 23	148.8	49.6	908	5	BU171410	AGENECOURT
C 24	148.6	49.5	345	8	AQ063008	CIT-HSP-2

ALIGNMENTS

RESULT 1  
AQ749749/c  
LOCUS  
DEFINITION HS\_5573 A2 C02 T7A RPCI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=1149 Col=4 Row=E, genomic survey sequence.  
ACCESSION AQ749749  
VERSION AQ749749.1 GI:5536907  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 828)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Kellier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
PUBMED 10449784  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 615-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.hsc.washington.edu  
Plate: 1149 row: E column: 4  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 828.  
Location/Qualifiers  
1..828  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate=1149 Col=4 Row=E"  
/sex="male"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

```
ORIGIN
Query Match      51.6%; Score 154.8; DB 8; Length 828;
Best Local Similarity 73.1%; Pred. No. 8.9e-20;
Matches 198; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
    |||||
Db 514 AAAATACAAAAATAGCTGGGCGTGTGTGTGCTGTATCCAGCTACTTGGGAGG 455
    |||||

QY 61 CTGAGGCGAGGAGCACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGACAGATAGTGC 120
    |||||
Db 454 CTGAGGCGAGGAGCAATCGCTGAACCCAGGAGGCGGAGTTGCAGTGACGCGGACCATGC 395
    |||||

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
    |||||
Db 394 CACTGCACCTCCAGCCTAGGTGACAGAAATGAGACTCTGTCTCAAAAAATAAAATA 335
    |||||

QY 181 AAAAATGCAGACTGTGATTCAGCAGCTCTGGGTTGAAGCCAGAACTCTCTGATAAATC 240
    |||||
Db 334 AAAAATGCATTAATGTTTTAAGACAGTTATGAATTTGTGTGGCGCCATTAGAGCCATC 275
    |||||

QY 241 AATGGCACTTAACTACTTGGAGGTCATGGAT 271
    |||||
Db 274 CTGGCTGCATCACTACCTACGCGGCGATGGAT 244
    |||||
```

RESULT 2  
BF871137 344 bp mRNA linear EST 17-JAN-2001  
LOCUS CM0-ET0121-311000-658-f04 ET0121 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF871137  
ACCESSION BF871137 GI:12261267  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldmann, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=CM0&t2=CM0-ET0121-311000-658-f04&t3=2000-10-31&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 26  
High quality sequence stop: 331.  
Location/Qualifiers  
1. .344  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source  
1. .526  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp459P1831"  
/tissue\_type="cortex"  
/dev\_stage="adult"  
/lab\_host="DH10B"

/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="ET0121"  
/notes="Organ: lung tumor; Vector: puc18; Site: 1: Smal;  
Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Query Match 50.3%; Score 151; DB 2; Length 344;  
Best Local Similarity 83.1%; Pred. No. 5.6e-19;  
Matches 172; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60  
 |||||  
Db 107 AGAATACAGAAATAGCCAGGCTGTGTGGCATGTGCTTAATCCCAAGCTACTCAGAGG 166  
 |||||  
QY 61 CTGAGGCGAGGAGCACTCTTGAATCCAGGAGGCGCAGGTTGCAGTGACGACAGATAGTGC 120  
 |||||  
Db 167 CTGAGGCGAGGAGAAATGGCGTGACCCAGGAGGCGGAGTTGCAGTGAGCCGAGATCGTC 226  
 |||||  
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180  
 |||||  
Db 227 CACTGCACCTCCAGCCTGGGCGACAGAGCAGACTCCGCTCTCAAAAAATAAAATA 286  
 |||||  
QY 181 AAAAATGCAGACTGTGATTCAGCAGGT 207  
 |||||  
Db 287 AAAAAGAAATTTTAAAAAATCCAGGT 313  
 |||||

RESULT 3  
CR556387/c 526 bp mRNA linear EST 12-JUL-2004  
LOCUS DKFZp459P1831\_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone  
DEFINITION DKFZp459P1831 5', mRNA sequence.  
ACCESSION CR556387.1 GI:50249989  
VERSION  
KEYWORDS EST.  
SOURCE Pongo pygmaeus (orangutan)  
ORGANISM Pongo pygmaeus

REFERENCE  
AUTHORS  
Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A., Fobo, G., Han, M. and Wiemann, S.  
Pongo pygmaeus mRNA (Koehler, K., Beyer, A., Mewes, H.W., et al.)  
Unpublished (2004)  
Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert from S. Wiemann,  
Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp459P1831) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clones@rzd.de  
Further information about the clone and the sequencing project is available at  
http://mips.gsf.de/projects/cdna/.  
Location/Qualifiers  
1. .526  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp459P1831"  
/tissue\_type="cortex"  
/dev\_stage="adult"  
/lab\_host="DH10B"

FEATURES  
source  
1. .526  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp459P1831"  
/tissue\_type="cortex"  
/dev\_stage="adult"  
/lab\_host="DH10B"

/clone lib="459 (synonym: pcor1)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiA; Site\_2: SfiB"

```

ORIGIN
Query Match      50.3%; Score 151; DB 7; Length 526;
Best Local Similarity 89.1%; Pred. No. 5.2e-19;
Matches 163; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCCGGAGG 60
    |||||
Db 387 AAATACAAAATAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCCGGAGG 328
    |||||

QY 61 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 120
    |||||
Db 327 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGATGAGATCACAC 268
    |||||

QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
    |||||
Db 267 CATTGCACTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 208
    |||||

QY 181 AAA 183
    |||
Db 207 AAA 205
    |||

```

```

RESULT 4
AQ750780
LOCUS
DEFINITION
  HS_5576_B2_E04_T7A_RPCI-11 Human Male BAC Library Homo sapiens
  genomic clone Plate=1152 Col=8 Row=J, genomic survey sequence.
ACCESSION
  AQ750780
VERSION
  AQ750780.1 GI:5537938
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 800)
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

```

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.html) or from Research Genetics (info@resgen.com). BAC end Web Server: <http://www.htbc.washington.edu>

```

Plate: 1152 row: J column: 8
Seq primer: T7
Class: BAC ends
High quality sequence stop: 800.
Location/Qualifiers
1. .800
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=1152 Col=8 Row=J"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor

```

and partially digested with a combination of EcoRI and EcoRII. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"

```

ORIGIN
Query Match      50.2%; Score 150.6; DB 8; Length 800;
Best Local Similarity 87.3%; Pred. No. 5.7e-19;
Matches 165; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCCGGAGG 60
    |||||
Db 59 AAATACAAAATAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCCGGAGG 118
    |||||

QY 61 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 120
    |||||
Db 119 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 178
    |||||

QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
    |||||
Db 179 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 238
    |||||

QY 181 AAAATGCA 189
    |||||
Db 239 AAAAAGTCA 247
    |||||

```

```

RESULT 5
AQ313572/c
AQ313572/c
LOCUS
DEFINITION
  RPCI11-101F17_TV_RPCI-11 Homo sapiens genomic clone RPCI-11-101F17,
  genomic survey sequence.
ACCESSION
  AQ313572
VERSION
  AQ313572.1 GI:4045035
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 675)
  Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
  Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  Use of human BAC End Sequences for Sequence-Ready Map Building
  Unpublished (1998)
Other GSSs: RPCI11-101F17.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbs@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .675
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7538536"
/db_xref="taxon:9606"
/clone="RPCI-11-101F17"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

```

FEATURES  
source

ORIGIN

```

Query Match      50.1%; Score 150.4; DB 8; Length 675;
Best Local Similarity 88.6%; Pred. No. 6.4e-19;
Matches 163; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
Db 343 AAAATACAAAATTAGCCAGGCGATGGTGGCATGTGCTGTAATCCAGCTACTTTGGGAGT 284

QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCGAGTTGTCAGTGACGAGATAGTGC 120
Db 283 CTGAGCAGGAGAACTCACCTGAATCCAGGAGCGCGAGTTGTCAGTGACGAGATAGTGC 224

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 223 CACTGCACCTCCAGCCTGGGTGACAAAGTGAGATTCTGTCTCAAAAAATAAAATAAATA 164

QY 181 AAAA 184
Db 163 AAAA 160

RESULT 6
BZ602060 757 bp DNA linear GSS 08-JUN-2003
LOCUS WHABV68TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-12K16, genomic survey sequence.
ACCESSION BZ602060
VERSION BZ602060.1 GI:31510522
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS Kowbel,D., Huang,G., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
source
1..757
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-12K16"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was
constructed from MCF7 breast_cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."
ORIGIN
Query Match      50.1%; Score 150.4; DB 8; Length 757;
Best Local Similarity 84.5%; Pred. No. 6.3e-19;
Matches 169; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
Db 205 AAAATACAAAATTAGCCAGGCGATGGTGGCATGTGCTGTCACAAAATAAAATAAATA 264

```

```

QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCGAGTTGTCAGTGACGAGATAGTGC 120
Db 265 CTGAAACGGGAGAACTCACTTGAATCCAGGAGCGGAGTTGTCAGTGACGCCAGATCGTGC 324

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 325 CACTGCACCTCCAGTGTGGGAGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 384

QY 181 AAAAATGCAGACTGTGATTC 200
Db 385 AAAGAGCCAGACAGTAAATC 404

RESULT 7
AQ474644/c
LOCUS CITBI-E1-2584M11.TR CITBI-E1 Homo sapiens genomic clone 2584M11,
DEFINITION genomic survey sequence.
ACCESSION AQ474644
VERSION AQ474644.1 GI:4652905
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 565)
AUTHORS Zhao,S., Adams,W.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
JOURNAL Contact: Shaying Zhao, William Nierman, Mark Adams
COMMENT Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1..565
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2584M11"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-E1"
/note="Vector: pBelobAC11; Site 1: EcoRI; Site 2: EcoRI;
Caltech Human BAC Library D"
ORIGIN
Query Match      50.1%; Score 150.2; DB 8; Length 565;
Best Local Similarity 87.7%; Pred. No. 7.2e-19;
Matches 164; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
Db 415 AAATGCAAAAAATTAGCCAGGCGGTGATGTCATGTCCTGTAGTCCAGCTACTTTGGGAGG 356

QY 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCGAGTTGTCAGTGACGAGATAGTGC 120
Db 355 CTGAGCAGGAGAACTGGCGTGAACCCGGGAGCGAGCTTGCAGTGAGCCGAGATAGCGC 296

QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 295 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 236

```

QY 181 AAAAAATG 187  
| | | | |  
Db 235 AAAAAAG 229

RESULT 8  
AQ490406/c  
LOCUS  
DEFINITION  
AQ490406 657 bp DNA linear GSS 24-APR-1999  
genomic survey sequence.  
ACCESSION  
AQ490406  
VERSION  
AQ490406.1 GI:4676280  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 657)  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and  
Venter,J.C.  
TITLE  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
JOURNAL  
Map Building  
COMMENT  
Unpublished (1997)  
Other\_GSSs: RPCI-11-244E4.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieterdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genet cs ([info@resgen.com](mailto:info@resgen.com)). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source  
1..657  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7593411"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-244E4"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 50.1%; Score 150.2; DB 8; Length 657;  
Best Local Similarity 74.9%; Pred. No. 7e-19;  
Matches 188; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1 ABAATATAAAATGACCGCGGTGATGTCATGCGCTGTAGTCCCGAGTACTCCGGAGG 60  
| | | | |  
Db 354 AAAAAACGAAAAATAGCCAGGCAATGTTGCACGTGCTGTATATCCAGCTACTCCGGAGG 295  
| | | | |  
QY 61 CTGAGCGAGGAGAACCTCTTTGAATCCAGGAGCGCGAGGTTCAGTGCAGCAGAGATAGTGC 120  
| | | | |  
Db 294 CCAGGCGAGGAGAAATCGCTTGAAACCCAGAGGTGGAGGTTCAGTGCAGTGCAGTGTGC 235  
| | | | |  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAATAAATA 180  
| | | | |  
Db 234 CACTGCACCTCCAGCTGGGTGACAGAGGAAGACTCTGTCTCAAAAAAATAAATAAATA 175  
| | | | |  
QY 181 AAAAAATGCAGACTGTGATTCAGCAGAGCTCTGGGTGAGGCCCAACTCTCTGATTAATTC 240  
| | | | |  
Db 174 AAAAAACAAAAAAGACAGCTTATGTTTATATATATTGGCAAGAAATAATAGAAAAATAA 115  
| | | | |

QY 241 AATGGCACTTA 251  
| | | | |  
Db 114 ATTGGAAATA 104

RESULT 9  
CF128305/c  
LOCUS  
DEFINITION  
CF128305 751 bp mRNA linear EST 05-AUG-2003  
UI-HE-ET0-awf-o-16-0-UI.r1 NIH\_MGC\_214 Homo sapiens cDNA clone  
IMAGE:30555903 5', mRNA sequence.  
ACCESSION  
CF128305  
VERSION  
CF128305.1 GI:33207416  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 751)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PubMed  
8889548  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Mary Hendrix  
CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/humanfl.html>  
The following repetitive elements were found in this cDNA  
sequence: 579-751, >ALU (matched complement)  
Seq primer: pYX-5.

FEATURES  
Location/Qualifiers  
1..751  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30555903"  
/tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_MGC\_214"  
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is GATNAGGCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 50.0%; Score 150; DB 7; Length 751;  
Best Local Similarity 86.4%; Pred. No. 7.5e-19;  
Matches 165; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 AAAAAATAAAATAGCCAGGCGGTGATGTCATGCGCTGTAGTCCAGCTACTCCGGAGG 60  
| | | | |  
Db 747 AAAAAACAAATNAGTCGCGTGTGGCGTGTGCTGTATCCAGCTACTCAGGAGG 688  
| | | | |  
QY 61 CTGAGCGAGGAGAACTCTTTGAATCCAGAGCGCAGGTTCAGTGCAGCAGAGATAGTGC 120  
| | | | |  
Db 687 CTGAGCGAGGAGAAATGCTTTGAACCTGGGAGCGGAGGTTCAGTGCAGTGCAGATAGTGC 628  
| | | | |

Tue Aug 9 17:01:28 2005

```

121  CACTGCACTCCAGCTGGGTGACAGAGTGTCTCTCAAAAAATAAAAAATA 180
122  |||||
123  627  CACTGATTCCAGCTGGGTGACAGAGTGTCTCTCAAAAAATAAAAAATA 568
124  |||||
125  181  AAAAAATGCAGA 191
126  |||||
127  567  AAAAGTCCAGA 557
128  |||||

RESULT 10
AW410354/c
LOCUS
DEFINITION
  fh05f06.y1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961538 3',
  mRNA sequence.
ACCESSION
  AW410354
VERSION
  EST.
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
REFERENCE
  1 (bases 1 to 491)
AUTHORS
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE
  NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Ling Hong/Rubin Laboratory
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: National Institutes of Health Intramural
  Sequencing Center (NIHC)
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/brrp/image/image.html
  Plate: LICM56 row: K column: 11
  Seq primer: M13RP1 reverse primer (ABI).
  Location/Qualifiers
    1. 491
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2961538"
      /tissue_type="rhabdomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_17"
      /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
      Site 2: XhoI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCAGCAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by
      Ling Hong in the laboratory of Gerald M. Rubin (University
      of California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
```

```

Query Match      49.9%; Score 149.8; DB 2; Length 491;
Best Local Similarity 88.1%; Pred. No. 8.9e-19;
Matches 163; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 AAAAAATAAAATAGCCAGGCGTGTCTCTGTGCTCTGTAGTCCAGCTACTCGGGAGG 60
DB 185 AAAATACAAAATATAGCCAGGCGTGTCTCTGTGCTCTGTAGTCCAGCTACTCGGGAGG 126
QY 61 CTGAGGAGGAGAACTCTTTGAATCCAGAGCGCGAGGTTGCAGTGACGAGATAGTC 120
DB 125 CTGAGGAGGAGAACTCTTTGAATCCAGAGCGCGAGGTTGCAGTGACGAGATAGTC 66
QY 121 CACTGCACTCCAGCTGGGTGACAGAGTGTCTCTCAAAAAATAAAAAATA 180
DB 65 CATTGTACTCCAGCTGGGCAACAGAGCGAGACTCTGTCTCAAAAAATAAAAAATA 6

FEATURES
source
  1. 572
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="RPIC1-11-73E12"
    /sex="Male"
    /cell_type="Lymphocytes"
    /clone_lib="RPIC1-11"
    /note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
    RPIC11 Human Male BAC Library"
```

```

Query Match      49.9%; Score 149.6; DB 8; Length 572;
Best Local Similarity 83.3%; Pred. No. 9.4e-19;
Matches 170; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1 AAAAAATAAAATAGCCAGGCGTGTCTCTGTGCTCTGTAGTCCAGCTACTCGGGAGG 60
DB 11 AAAATACAAAATATAGCCAGGCGTGTCTCTGTGCTCTGTAGTCCAGCTACTCGGGAGG 70
QY 61 CTGAGGAGGAGAACTCTTTGAATCCAGAGCGCGAGGTTGCAGTGACGAGATAGTC 120
DB 71 CTGAGGAGGAGAACTCTTTGAATCCAGGCGGAGGTTACAGTCAGCCGAGATAGTC 130
QY 121 CACTGCACTCCAGCTGGGTGACAGAGTGTCTCTCAAAAAATAAAAAATA 180
DB 131 CACTGCACTCCAGCTGGGTGACAGAGTGTCTCTCAAAAAATAAAAAATA 190
QY 181 AAAATGCGAGCTGTCTCTCAAGCA 204
DB 191 GAAAAAGAAAAGAGAAAGCATCA 214

RESULT 12
```

```

QY 181 AAAAA 185
DB 5 AAAAA 1

RESULT 11
AQ268868
LOCUS
DEFINITION
  AQ268868 572 bp DNA linear GSS 27-APR-1999
  genomic survey sequence.
ACCESSION
  AQ268868
VERSION
  GSS.
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
REFERENCE
  1 (bases 1 to 572)
AUTHORS
  Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
  Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
  Use of human BAC End Sequences for Sequence-Ready Map Building
  Unpublished (1998)
TITLE
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0200
  Fax: 301 838 0208
  Email: mdadams@tigr.org
  Clones are derived from the human BAC library RPIC1-11. For BAC
  library availability, please contact Pieter de Jong
  (pieter@dejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
  Research Genetics (info@resgen.com). BAC end search page:
  http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
  Seq primer: SP6
  Class: BAC ends.
  Location/Qualifiers
    1. 572
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="RPIC1-11-73E12"
      /sex="Male"
      /cell_type="Lymphocytes"
      /clone_lib="RPIC1-11"
      /note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
      RPIC11 Human Male BAC Library"
```

AQ312235  
LOCUS  
RPC111-103E16.TJ RPCI-11 linear GSS 04-MAY-1999  
genomic survey sequence.  
ACCESSION  
AQ312235  
VERSION  
AQ312235.1 GI:4043899  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 638)  
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.  
Use of human BAC End Sequences for Sequence-Ready Map Building  
Unpublished (1998)  
TITLE  
Other\_GSSs: RPC111-103E16.TV  
JOURNAL  
COMMENT  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeetigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1..638  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7539279"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-103E16"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/notes="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPC111 Human Male BAC Library"

ORIGIN  
Query Match 49.9%; Score 149.6; DB 8; Length 638;  
Best Local Similarity 83.3%; Pred. No. 9.2e-19;  
Matches 170; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
QY 1 AAAATATAAAATTAGCCAGGCGTGATGTGCATGTGCTGTAGTCCAGCTACTTCGGGAGG 60  
|||||  
Db 56 AAATACAAAATTAGCCGGGCGATGTGGCATGTGCTGTAGTCCAGCTACTTCGGGAGG 115  
|||||  
QY 61 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGGCGCAGGTTGCAGTGACGAGATAGTGC 120  
|||||  
Db 116 CTGAGGCGGAGAAATTTGCTTGAACCCGGAGGCGGAGGTTACAGTCAGCGGAGATGTC 175  
|||||  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180  
|||||  
Db 176 CACTGCACCTCCAGCTGGGTGATAGAGCGAGACTCTGTCTCAAAAAATAAAATAAATA 235  
|||||  
QY 181 AAAAATGCAGACTGTGATTCAGCA 204  
|||||  
Db 236 GAAAAAGAAAGAAAGGCAATCA 259  
|||||

RESULT 13  
AG047289  
LOCUS  
AG047289 642 bp DNA linear GSS 02-NOV-2001  
DEFINITION  
Pan troglodytes DNA, clone: PTB-026M07.F, genomic survey sequence.  
ACCESSION  
AG047289  
VERSION  
AG047289.1 GI:16584181

GSS.  
Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE  
1  
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
BAC end sequences of Library PTB  
Unpublished  
2 (bases 1 to 642)  
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chimpes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1..642  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-026M07.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"

ORIGIN  
Query Match 49.8%; Score 149.4; DB 9; Length 642;  
Best Local Similarity 84.4%; Pred. No. 1e-18;  
Matches 168; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
QY 1 AAAATATAAAATTAGCCAGGCGTGATGTGCATGTGCTGTAGTCCAGCTACTTCGGGAGG 60  
|||||  
Db 106 AAATACAAAATTAGCCAGGCGATGTGTGCATGCCTATATATCCAGCTACTCAGAGA 165  
|||||  
QY 61 CTGAGGCGAGGAGAACCTCTTGAATCCAGGAGGCGCAGGTTGCAGTGACGAGATAGTGC 120  
|||||  
Db 166 CTGAGGCGAGGAGAACCTGTCTGAACCCAGGAGGAGGTTGCAGTGAGCTGGGATAGTC 225  
|||||  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180  
|||||  
Db 226 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAAGG 285  
|||||  
QY 181 AAAAATGCAGACTGTGATT 199  
|||||  
Db 286 AAAGTAGAAAATTGACT 304  
|||||

RESULT 14  
AG160213  
LOCUS  
AG160213 677 bp DNA linear GSS 09-JAN-2002  
DEFINITION  
Pan troglodytes DNA, clone: RP43-025P15.T7, genomic survey  
sequence.  
ACCESSION  
AG160213  
VERSION  
AG160213.1 GI:16689891  
KEYWORDS  
GSS.  
SOURCE  
Pan troglodytes (chimpanzee)  
ORGANISM  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE  
1

Tue Aug 9 17:01:28 2005

**AUTHORS**  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Tokui, Y., Watanabe, H. and Sakaki, Y.  
**TITLE**  
BAC end sequences of Library RPCI-43  
**JOURNAL**  
Unpublished  
**REFERENCE**  
**AUTHORS**  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Tokui, Y., Watanabe, H. and Sakaki, Y.  
**TITLE**  
Direct Submission  
**JOURNAL**  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
[E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170]  
**COMMENT**  
Clones are derived from the chimpanzee BAC library RPCI-43. This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
**PRIMERS**  
Sequencing: T7  
**LIBRARY**  
Vector : pBACe3.6  
R-site 1 : EcoRI  
R-site 2 : EcoRI  
**FEATURES**  
source  
Location/Qualifiers  
1. .677  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-025P15.T7"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RPCI-43 Chimpanzee Male BAC Library"  
**ORIGIN**  
Query Match 49.8%; Score 149.4; DB 9; Length 677;  
Best Local Similarity 84.4%; Pred. No. 9.9e-19;  
Matches 168; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
QY 1 AAAATATATAAAATTAGCCAGCGGTGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
Db 32 AAAATTCAAACATTAGCCAGCGGTGTCATGTGCTGTAGTCCAGCTACTCGGAGG 91  
QY 61 CTGAGGAGGACCTCTTGAATCCAGGAGCGGAGTTCAGTGCAGAGATAGTGC 120  
Db 92 CTGAGACAGGAGATCACTTGAACCCAGGAGCGGGGTTCAGTGCAGATCTCAC 151  
QY 121 CACTGCACTCCAGCGTGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAATA 180  
Db 152 CACTGCATTCCAGCGTGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAATA 211  
QY 181 AAAATGCGAGCTGTGATT 199  
Db 212 AAAAAAAGCGCTTTTAAT 230  
**RESULT 15**  
AQ890095/c 732 bp DNA linear GSS 10-NOV-1999  
LOCUS HS 3100 Al\_E11 T7C CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=3100 Col=21 Row=I, genomic survey  
sequence.  
ACCESSION AQ890095.1 GI:6346285  
VERSION AQ890095  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 732)  
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome

**JOURNAL**  
MEDLINE  
PUBMED  
**COMMENT**  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
1049764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end web server: http://www.htsc.washington.edu  
Plate: 3100 row: I column: 21  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 732.  
**FEATURES**  
source  
Location/Qualifiers  
1. .732  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=3100 Col=21 Row=I"  
/sex="male"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in  
E-Coli DH10B"  
**ORIGIN**  
Query Match 49.8%; Score 149.4; DB 8; Length 732;  
Best Local Similarity 79.4%; Pred. No. 9.8e-19;  
Matches 177; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
QY 1 AAAATATATAAAATTAGCCAGCGGTGTCATGTGCTGTAGTCCAGCTACTCGGAGG 60  
Db 541 AAACACAAAAATTAGCCAGCGGTGTCATGTGCTGTAGTCCAGCTACTCGGAGG 482  
QY 61 CTGAGGAGGAGAACCTTTCAATCCAGAGCGGAGGTTCCAGTGACGAGATAGTGC 120  
Db 481 CTGAGGAGGAGAAATTGCTTGAATCCAGGAGCGGAGGTTCCAGTGAGCTGAGATCGTAC 422  
QY 121 CACTGCATCTCCAGCGTGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAATA 180  
Db 421 TACTGCATCTCCAGCGTGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAATA 362  
QY 181 AAAAATGCGAGCTGTGATTTCAGCAGGCTCGGGTTGAAGCCAG 223  
Db 361 AAACAAACAAACAAACAAACAAACAAATTTGAATGATGCCAG 319  
Search completed: August 4, 2005, 14:32:36  
Job time : 1422.95 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 61.8219 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-6  
Perfect score: 300  
Sequence: 1 aaatataaaattagccag.....aatctaataagaagtactga 300

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	156.4	52.1	24496	4	US-09-949-016-11823 Sequence 11823, A
2	156.4	52.1	24497	4	US-09-949-016-14253 Sequence 14253, A
3	154.4	51.5	23155	4	US-09-949-016-12366 Sequence 12366, A
C 4	154	51.3	601	4	US-09-949-016-12366 Sequence 80246, A
C 5	154	51.3	68173	4	US-09-949-016-14046 Sequence 14046, A
C 6	153.4	51.1	601	4	US-09-949-016-78059 Sequence 78059, A
C 7	153.4	51.1	36759	4	US-09-949-016-12216 Sequence 12216, A
C 8	153.4	51.1	36760	4	US-09-949-016-14021 Sequence 14021, A
C 9	153	51.0	601	4	US-09-949-016-204885 Sequence 204885, A
10	153	51.0	30221	4	US-09-949-016-12577 Sequence 12577, A
11	153	51.0	30222	4	US-09-949-016-17299 Sequence 17299, A
C 12	153	51.0	128723	4	US-09-949-016-17533 Sequence 17533, A
C 13	152.6	50.9	601	4	US-09-949-016-204886 Sequence 204886, A
C 14	152.6	50.9	601	4	US-09-949-016-204887 Sequence 204887, A
15	152	50.7	601	4	US-09-949-016-143301 Sequence 143301, A
16	152	50.7	601	4	US-09-949-016-143472 Sequence 143472, A
17	151.8	50.6	421491	4	US-09-949-016-12805 Sequence 12805, A
18	151.8	50.6	421494	4	US-09-949-016-14060 Sequence 14060, A
C 19	151.4	50.5	22539	4	US-09-949-016-15931 Sequence 15931, A
C 20	151	50.3	40435	4	US-09-949-016-16504 Sequence 16504, A
C 21	150.8	50.3	120727	4	US-09-949-016-15787 Sequence 15787, A
C 22	150.8	50.3	120727	4	US-09-949-016-15788 Sequence 15788, A
C 23	150.8	50.3	240157	4	US-09-949-016-16264 Sequence 16264, A
24	150.6	50.2	601	4	US-09-949-016-38743 Sequence 38743, A
25	150.6	50.2	601	4	US-09-949-016-143733 Sequence 143733, A
26	150.6	50.2	601	4	US-09-949-016-177920 Sequence 177920, A
C 27	150.6	50.2	34725	4	US-09-949-016-15797 Sequence 15797, A

C 28	150.6	50.2	34765	4	US-09-949-016-13808 Sequence 13808, A
C 29	150.6	50.2	86936	4	US-09-949-016-17314 Sequence 17314, A
C 30	150.4	50.1	601	4	US-09-949-016-115543 Sequence 115543, A
C 31	150.4	50.1	601	4	US-09-949-016-115636 Sequence 115636, A
C 32	150.4	50.1	601	4	US-09-949-016-115729 Sequence 115729, A
C 33	150.4	50.1	601	4	US-09-949-016-115822 Sequence 115822, A
C 34	150.4	50.1	601	4	US-09-949-016-115915 Sequence 115915, A
C 35	150.4	50.1	601	4	US-09-949-016-116008 Sequence 116008, A
C 36	150.4	50.1	601	4	US-09-949-016-116101 Sequence 116101, A
C 37	150.4	50.1	601	4	US-09-949-016-116194 Sequence 116194, A
C 38	150.4	50.1	601	4	US-09-949-016-116287 Sequence 116287, A
C 39	150.4	50.1	601	4	US-09-949-016-116380 Sequence 116380, A
C 40	150.4	50.1	601	4	US-09-949-016-143300 Sequence 143300, A
C 41	150.4	50.1	601	4	US-09-949-016-143471 Sequence 143471, A
C 42	150.4	50.1	601	4	US-09-949-016-150076 Sequence 150076, A
C 43	150.4	50.1	26115	4	US-09-949-016-15959 Sequence 15959, A
C 44	150.4	50.1	37792	4	US-09-949-016-14503 Sequence 14503, A
C 45	150.4	50.1	37795	4	US-09-949-016-14263 Sequence 14263, A

ALIGNMENTS

RESULT 1

US-09-949-016-11823  
; Sequence 11823, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11823  
; LENGTH: 24496  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11823

Query Match	52.1%	Score	156.4	DB	4	Length	24496
Best Local Similarity	85.0%	Pred. No.	2.1e-38				
Matches	175	Conservative	0	Mismatches	31	Indels	0
Gaps	0						
QY	1	AAAAATATAAATTAGCAGCGGTGATGTCATGCTGCTGTAGTCCAGCTACTCGGGAGG	60				
Db	8078	AAATATACAGAAATTAGCTGGGTGTGATGCTGCTGCTGTAATCCAGCTACTCGGGAGC	8137				
QY	61	CTGAGCAGGAGAACTCTTGAATCCAGAGCGCGAGTTGCGAGTCAGAGATAGTGC	120				
Db	8138	CTGAGCAGGAGGAACTCTTGAATCCGGAGCGGAGTTGCGAGCGGAGATTGTAC	8197				
QY	121	CACTGCACCTCCAGCGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAATAAATA	180				
Db	8198	CACTGCACCTCCAGCGCTGGGCCACAGAGAAGACTCTGTCTCAAAAAATAAATAAATA	8257				
QY	181	AAAAATCGAGCTGTGATTTCAGCAGG	206				
Db	8258	AAAAATCCAGCTGAGTCTCACCAGG	8283				
RESULT 2							
US-09-949-016-14253							
; Sequence 14253, Application US/09949016							
; Patent No. 6812339							

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14253
; LENGTH: 24497
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14253

Query Match 52.1%; Score 156.4; DB 4; Length 24497;
Best Local Similarity 85.0%; Pred. No. 2.1e-38;
Matches 175; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
DB 8078 AAAATACAGAAATAGCTGGGTGTGATGTCGTGCTGTAAATCCAGCTACTCGGAGC 8137

QY 61 CTGAGGCGAGAGACTCTTGAATCCAGGAGCGCAGGTTGCAGTCAGCAGATAGTGC 120
DB 8138 CTGAGGCGAGGGAATCTTTGAATCCGGAGCGGAGTTGCAGTCAGCGGAGTTGTAC 8197

QY 121 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
DB 8198 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATAA 8257

QY 181 AAAATGCAGCTGTGATTCAGCAGGTCCTGGTTGAAGCC 206
DB 8258 AAAATCCAGCTGAGTCTCACCAGG 8283

RESULT 3
US-09-949-016-12366
; Sequence 12366, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12366
; LENGTH: 23155
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) - (23155)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12366

Query Match 51.5%; Score 154.4; DB 4; Length 23155;
Best Local Similarity 81.4%; Pred. No. 8.8e-38;

```

```

Matches 179; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
DB 8807 AAAATATAAAATAGCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 8866

QY 61 CTGAGGCGAGAGACTCTTGAATCCAGGAGCGCAGGTTGCAGTCAGCAGATAGTGC 120
DB 8867 CTGAGGCGAGGAGATCACTTGAACCCGGAGCGCAGGTTGCAGTCAGCTGAGATTGCAC 8926

QY 121 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
DB 8927 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATAA 8986

QY 181 AAAATGCAGCTGTGATTCAGCAGGTCCTGGTTGAAGCC 220
DB 8987 ACAAGTGGGACTGAACTTCCCAACCTCTCCTTGTGTCTC 9026

RESULT 4
US-09-949-016-80246/c
; Sequence 80246, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80246
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-80246

Query Match 51.3%; Score 154; DB 4; Length 601;
Best Local Similarity 81.7%; Pred. No. 2e-38;
Matches 178; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGAGG 60
DB 491 AAAATACAAAAAGTAGCCAGCGGTGTCATGTCCTGTAGTCCAGCTACTCGGAGG 432

QY 61 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTCAGCAGATAGTGC 120
DB 431 CTGAGGCGAGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTCAGCTGAGATTGCAC 372

QY 121 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
DB 371 CACTGCCTCCAGCTGGGTGCAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATAA 312

QY 181 AAAATGCAGCTGTGATTCAGCAGGTCCTGGTTGAAG 218
DB 311 AAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 274

RESULT 5
US-09-949-016-14046/c
; Sequence 14046, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

Db	14	AAATAGAAAAATTAGCCGGCGATGTGGCATGCACCTGTAGTCCAGACTACTCTGGGAGG	73
Qy	61	CTGAGGAGGAGAACTCTTTGAATCCAGGAGCGCAGGTTGCAGTGAGAGATAGTGC	120
Db	74	CTGAGGAGGAGGATCACTTGAACCCAGGAGCGGAGGTTGCAGTGAGAGATCGGC	133
Qy	121	CATGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA	180
Db	134	CATGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA	193
Qy	181	AAAAATCCAGACTGTGATTTCAGC	203
Db	194	AGAAATAACAGAAAGCTTCATC	216
<p>RESULT 7</p> <p>US-09-949-016-12216/c</p> <p>; Sequence 12216, Application US/09949016</p> <p>; Patent No. 6812339</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: VENTER, J. Craig et al.</p> <p>; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED</p> <p>; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF</p> <p>; FILE REFERENCE: CL001307</p> <p>; CURRENT APPLICATION NUMBER: US/09/949,016</p> <p>; CURRENT FILING DATE: 2000-04-14</p> <p>; PRIOR APPLICATION NUMBER: 60/241,755</p> <p>; PRIOR FILING DATE: 2000-10-20</p> <p>; PRIOR APPLICATION NUMBER: 60/237,768</p> <p>; PRIOR FILING DATE: 2000-10-03</p> <p>; PRIOR APPLICATION NUMBER: 60/231,498</p> <p>; PRIOR FILING DATE: 2000-09-08</p> <p>; NUMBER OF SEQ ID NOS: 207012</p> <p>; SOFTWARE: Fast-SEQ for Windows Version 4.0</p> <p>; SEQ ID NO 12216</p> <p>; LENGTH: 36759</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Human</p> <p>US-09-949-016-12216</p>			
<p>Query Match 51.1%; Score 153.4; DB 4; Length 36759;</p> <p>Best Local Similarity 84.7%; Pred. No. 2.3e-37;</p> <p>Matches 172; Conservative 0; Mismatches 31; Indels 0; Gaps 0;</p>			
Qy	1	AAAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGACTACTCTGGGAGG	60
Db	14576	AAATAGAAAAATTAGCCGGCATGTGGCATGCACCTGTAGTCCAGACTACTCTGGGAGG	14517
Qy	61	CTGAGGAGGAGAACTCTTTGAATCCAGGAGCGCAGGTTGCAGTGAGAGATAGTGC	120
Db	14516	CTGAGGAGGAGGATCACITTTGAACCCAGGAGCGGAGGTTGCAGTGAGAGATCGGC	14457
Qy	121	CATGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA	180
Db	14456	CATGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA	14397
Qy	181	AAAAATCCAGACTGTGATTTCAGC	203
Db	14396	AGAAATAACAGAAAGCTTCATC	14374
<p>RESULT 8</p> <p>US-09-949-016-14021/c</p> <p>; Sequence 14021, Application US/09949016</p> <p>; Patent No. 6812339</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: VENTER, J. Craig et al.</p> <p>; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED</p> <p>; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF</p> <p>; FILE REFERENCE: CL001307</p> <p>; CURRENT APPLICATION NUMBER: US/09/949,016</p> <p>; CURRENT FILING DATE: 2000-04-14</p> <p>; PRIOR APPLICATION NUMBER: 60/241,755</p>			
<p>Query Match 51.1%; Score 153.4; DB 4; Length 601;</p> <p>Best Local Similarity 84.7%; Pred. No. 3e-38;</p> <p>Matches 172; Conservative 0; Mismatches 31; Indels 0; Gaps 0;</p>			
Qy	1	AAAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGACTACTCTGGGAGG	60

Db 439 CACTGCACCTCCAGCCTGGATGACAGAGCGAGACTCTGTCTCCAAAAA 380  
Qy 181 AAAAA 185  
Db 379 AAAAA 375  
RESULT 10  
US-09-949-016-12577  
; Sequence 12577, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12577  
; LENGTH: 30221  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12577  
Query Match 51.0%; Score 153; DB 4; Length 30221;  
Best Local Similarity 81.8%; Pred. No. 2.7e-37;  
Matches 189; Conservative 0; Mismatches 40; Indels 2; Gaps 1;  
Qy 1 AAAAAATTAAGCCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 60  
Db 11569 AAAAAATTAAGCCAGCGGTGTCGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 11628  
Qy 61 CTGAGCGAGGAGAACTCTTGAATCCAGGAGCGCAGGTTCGAGTGACGAGATAGTGC 120  
Db 11629 CTGTGGCATGAGAACTCTTGAACAGGAGGATGATGTTGAGTGAGATTTGTC 11688  
Qy 121 CACTGCACCTCCAGCCTGGGTGACAGTGACAGTCTGTCTCAAAAAATAAATA 180  
Db 11689 CACTGAACTGCAAGCTGGGTGACAGCAAGACTCTGTCTCAAAAAATAAATA 11748  
Qy 181 AAAAA--TGCACACTGTGATTCAGCAGGTCTGGGTTGAAGCCAGAACTCT 229  
Db 11749 AAAGACCTGCAGATCTGATTCAGTGGTTCGGGGTGAGGCCAAGAGTCT 11799  
RESULT 11  
US-09-949-016-17299  
; Sequence 17299, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17299

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14021  
; LENGTH: 36760  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14021  
Query Match 51.1%; Score 153.4; DB 4; Length 36760;  
Best Local Similarity 84.7%; Pred. No. 2.3e-37;  
Matches 172; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
Qy 1 AAAAAATTAAGCCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 60  
Db 14576 AAAAAATTAAGCCAGCGGTGTCGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 14517  
Qy 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCAGGTTCGAGTGACGAGATAGTGC 120  
Db 14516 CTGAGCAGGAGATCACTTGAACCCAGGAGCGGAGTTGAGTGAGCGGATCGGC 14457  
Qy 121 CACTGCACCTCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAAATAAATA 180  
Db 14456 CACTGCACCTCCAGCCTGGGTGACAGAGCAAGACTCTGTCTCAAAAAATAAATA 14397  
Qy 181 AAAAAATTAAGCCAGCTGTGATTCAGC 203  
Db 14396 AAAAAATTAAGCAAGAGCTTCATC 14374  
RESULT 9  
US-09-949-016-204885/c  
; Sequence 204885, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 204885  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-204885  
Query Match 51.0%; Score 153; DB 4; Length 601;  
Best Local Similarity 89.2%; Pred. No. 4e-38;  
Matches 165; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
Qy 1 AAAAAATTAAGCCAGCGGTGATGTCATGTCCTGTAGTCCAGCTACTCGGGAGG 60  
Db 559 AAAAAATTAAGCCAGCGGTGATGTCGCGGACCTGTAGTCCAGCTACTCGGGAGG 500  
Qy 61 CTGAGCAGGAGAACTCTTGAATCCAGGAGCGCAGGTTCGAGTGACGAGATAGTGC 120  
Db 499 CTGAGCATGAGAACTCTTGAACCCAGGAGCGGAGTTGAGTGAGCCGAGATCGAGC 440  
Qy 121 CACTGCACCTCCAGCCTGGGTGACAGTGAGACTCTGTCTCAAAAAATAAATA 180

;  
; LENGTH: 30222  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17299

Query Match 51.0%; Score 153; DB 4; Length 30222;  
Best Local Similarity 81.8%; Pred. No. 2.7e-37;  
Matches 189; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 60  
|||||  
Db 11569 AAATACAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 11628  
|||||  
QY 61 CTGAGCAGGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGAGATAGTGC 120  
|||||  
Db 11629 CTGTGCATGAGAATCTTTGAACCCAGGAGGTTGATGTTGCAGTGCAGTGCAGTGC 11688  
|||||  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180  
|||||  
Db 11689 CACTGAACCTGCAGCTGGGTGACAGAGCAAGACTCTGTCTCAAAAAATAAAATAA 11748  
|||||  
QY 181 AAAAA--TGCACTCTGATTACAGCAGGCTGGGTTGAAGCCAGCACTCT 229  
|||||  
Db 11749 AAGACCTGCAGATCCTGATTACAGTGTCTGGGTGAGGCCCAAGAGTCT 11799  
|||||

## RESULT 12

US-09-949-016-17533/c  
; Sequence 17533, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17533

; LENGTH: 128723

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17533

Query Match 51.0%; Score 153; DB 4; Length 128723;  
Best Local Similarity 89.2%; Pred. No. 5.6e-37;  
Matches 165; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 60  
|||||  
Db 113491 AAATACAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 113432  
|||||  
QY 61 CTGAGCAGGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGAGATAGTGC 120  
|||||  
Db 113431 CTGAGGCGATGAGAATCACTTTGAACCCAGGAGCGAGGTTGCAGTGCAGCAGATCGAGC 113372  
|||||  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180  
|||||  
Db 113371 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAA 113312  
|||||  
QY 181 AAAAA 185  
|||||  
Db 113311 AAAAA 113307  
|||||

## RESULT 13

Query Match 50.9%; Score 152.6; DB 4; Length 601;  
Best Local Similarity 88.6%; Pred. No. 5.4e-38;  
Matches 164; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

US-09-949-016-204886/c  
; Sequence 204886, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 204886

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-204886

Query Match 50.9%; Score 152.6; DB 4; Length 601;  
Best Local Similarity 88.6%; Pred. No. 5.4e-38;  
Matches 164; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 60  
|||||  
Db 422 AAATACAAAATTAGCCAGGCGTGATGTCATGCTGTAGTCCAGCTACTCGGGAGG 363  
|||||  
QY 61 CTGAGCAGGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGAGATAGTGC 120  
|||||  
Db 362 CTGAGGCGATGAGAATCACTTTGAACCCAGGAGCGAGGTTGCAGTGCAGCAGATCGAGC 303  
|||||  
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180  
|||||  
Db 302 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 243  
|||||  
QY 181 AAAAA 185  
|||||  
Db 242 AAAAA 238  
|||||

## RESULT 14

US-09-949-016-204887/c  
; Sequence 204887, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 204887

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-204887

Query Match 50.9%; Score 152.6; DB 4; Length 601;  
Best Local Similarity 88.6%; Pred. No. 5.4e-38;  
Matches 164; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

Qy	1	AAAAATAAAAAATTAGCCAGCGGTGATGTCAATGTGCTCTAGTCCCAAGCTACTCTCGGAGG	60
Db	326	AAAAATCAAAAAATTAGCCAGGCGTGTGCGGGGCACTGTAGTCCCAAGCTACTCTCGGAGG	267
Qy	61	CTGAGCCAGGAGAACTCTTGAATCCAGGAGCGCAGGTTGCAGTGTAGCAGAGATAGTGC	120
Db	266	CTGAGGCATGAGAACTCACTTTGAACCCAGAGGCGGAGGTTGCAGTGTAGCCGAGATCGAGC	207
Qy	121	CAGTGCACCTCCAGCGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAAATAAAAAATA	180
Db	206	CAGTGCACCTCCAGCGCTGGATGACAGAGCGAGACTCTGTCTCAAAAAAATAAAAAA	147
Qy	181	AAAAA	185
Db	146	AAAAA	142

RESULT 15

```

US-09-949-016-143301
; Sequence 143301, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143301
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-143301

```

	Query Match	50.7%;	Score 152;	DB 4;	Length 601;
	Best Local Similarity	88.2%;	Pred. No. 8.3e-38;		
	Matches 164;	Conservative	1;	Mismatches -21;	Indels 0;
					Gaps 0;
Qy	1	AAAAATATAAAATTTAGCCAGGCGGTGATGTCATGTGCGCTGTAGTCCAGCTACTCGGGAGG	60		
Db	208	AAAAATATAAAATTTAGCTGGGCGTGGTGGCGGGCGCTGTAGTCCAGCTACTCGGGAGG	267		
Qy	61	CTGAGCGAGGAAACCTTTTGAATCCAGAGCGCGAGTTGCAGTGACGACGATAGTGC	120		
Db	268	CTGAGCGAGGAAATTGCTTTGAACCCAGGAGCGAGGTTGCAGTGAGCCGGAGATCGCAC	327		
Qy	121	CAGTCGACTCCAGCGCTGGGTGACAGAGTGAGAGCTCTGTCTCAAAAAATAATAAAATA	180		
Db	328	CAGTCGACTCCAGCGCTGGGTGACAGAGTGAGAGCTCCGCCCTCAAAAAATAAAAAATA	387		
Qy	181	AAAAAT	186		
Db	388	AAAAAT	393		

Search completed: August 4, 2005, 14:40:58  
Job time : 64.8219 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 535.142 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-6  
Perfect score: 300  
Sequence: 1 aaatatataaaattgacag.....aatctaataagaagctactga 300

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues  
Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	US-09-899-276-6	Sequence 6, Appli
2	300	100.0	11793	US-10-685-705-4	Sequence 4, Appli
3	167.8	55.9	5926	US-10-311-455-1625	Sequence 1625, Ap
4	160.8	53.6	31926	US-10-322-281-712	Sequence 712, App
c 5	156.6	52.2	552	US-10-027-632-128663	Sequence 128663,
c 6	156.6	52.2	552	US-10-027-632-128663	Sequence 128663,
c 7	156.4	52.1	5304	US-09-764-847-1214	Sequence 1214, Ap

RESULT 1  
US-09-899-276-6  
; Sequence 6, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Dellus, Hajjo  
; APPLICANT: Pousetka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899, 276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-6

ALIGNMENTS

Sequence 1215, Ap  
Sequence 1214, Ap  
Sequence 1215, Ap  
Sequence 128662,  
Sequence 128662,  
Sequence 3426, Ap  
Sequence 128664,  
Sequence 128664,  
Sequence 66, Appl  
Sequence 135370,  
Sequence 135371,  
Sequence 135370,  
Sequence 135371,  
Sequence 257834,  
Sequence 257834,  
Sequence 1463, Ap  
Sequence 1576, Ap  
Sequence 50, Appl  
Sequence 3428, Ap  
Sequence 17645, A  
Sequence 5, Appli  
Sequence 8827, Ap  
Sequence 8827, Ap  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 58, Appl  
Sequence 116534,  
Sequence 116534,  
Sequence 17556, A  
Sequence 1666, Ap  
Sequence 1666, Ap  
Sequence 2650, Ap  
Sequence 248, App  
Sequence 98423, A  
Sequence 98424, A  
Sequence 98423, A  
Sequence 98424, A

Query Match 100.08; Score 300; DB 9; Length 300;  
Best Local Similarity 100.08; Pred. No. 8.4e-82;  
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAAATATATAAAATTAGCCAGCGGTGATGTGTCATGTCCTGTAGTCCACGCTACTCGGGAGG 60
Db 1 AAAATATATAAAATTAGCCAGCGGTGATGTGTCATGTCCTGTAGTCCACGCTACTCGGGAGG 60
QY 61 CTGAGGCGAGGAGAACTCTTGAATCCAGAGCGCGAGGTTGAGTGAGCAGAGATAGTGC 120
Db 61 CTGAGGCGAGGAGAACTCTTGAATCCAGAGCGCGAGGTTGAGTGAGCAGAGATAGTGC 120
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
QY 181 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGGTTGAAGCCCAAGAACTCTCTGATAAATC 240
Db 181 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGGTTGAAGCCCAAGAACTCTCTGATAAATC 240
QY 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTAATCTAATAGAGCTACTGA 300
Db 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTAATCTAATAGAGCTACTGA 300

RESULT 2
US-10-685-705-4
; Sequence 4, Application US/10685705
; Publication No. US2004017387A1
; GENERAL INFORMATION:
; APPLICANT: University of Kentucky Research Foundation
; APPLICANT: JAYAKRISHNA, Ambati
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular
; TITLE OF INVENTION: Degeneration
; FILE REFERENCES: 050229-0415
; CURRENT APPLICATION NUMBER: US/10/685,705
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: 60/422,096
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-685-705-4

Query Match 100.0%; Score 300; DB 19; Length 11793;
Best Local Similarity 100.0%; Pred. No. 4.1e-81;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAATATATAAAATTAGCCAGCGGTGATGTGTCATGTCCTGTAGTCCACGCTACTCGGGAGG 60
Db 2500 AAAATATATAAAATTAGCCAGCGGTGATGTGTCATGTCCTGTAGTCCACGCTACTCGGGAGG 2559
QY 61 CTGAGGCGAGGAGAACTCTTGAATCCAGAGCGCGAGGTTGAGTGAGCAGAGATAGTGC 120
Db 2560 CTGAGGCGAGGAGAACTCTTGAATCCAGAGCGCGAGGTTGAGTGAGCAGAGATAGTGC 2619
QY 121 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 2620 CACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 2679
QY 181 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGGTTGAAGCCCAAGAACTCTCTGATAAATC 240
Db 2680 AAAAATGCAGACTGTGATTCAGCAGGTCCTGGGTTGAAGCCCAAGAACTCTCTGATAAATC 2739
QY 241 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTAATCTAATAGAGCTACTGA 300
Db 2740 AATGGCACTTAACCTACTTGGAGGTCATGGATGCTTTGCTAATCTAATAGAGCTACTGA 2799

RESULT 3
US-10-311-455-1625
; Sequence 1625, Application US/10311455
; Publication No. US20030143606A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1625
; LENGTH: 5926
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1625

Query Match 55.9%; Score 167.8; DB 15; Length 5926;
Best Local Similarity 80.7%; Pred. No. 1.3e-40;
Matches 196; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 58 AGGCTGAGGCGAGGAGAACCTCTTGAATCCAGAGCGCGAGGTTGAGTGAGCAGAGATAG 117
Db 1 AGGTTGAGGTAGGAGAAATTTTGAATTTAGGAGCGGTAGGTTGAGTGAGCAGAGATAG 60
QY 118 TGCCACTGCACCTCCAGCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAA 177
Db 61 TGTTATTGTTATTAGTTGGGTGATAGAGTGAGATTTGTTTTAAAAATAAAATAAATA 120
QY 178 ATAAAAATGCAGACTGTGATTTCAGCAGGTCGGTTGAAGCCCAAGAACTCTCTGATAA 237
Db 121 ATAAAAATGCAGACTGTGATTTCAGCAGGTTGGGTTGAAGCTTTAGAAATTTTGTATA 180
QY 238 TTCAATGGCAGCTTAACCTACTTGGAGGTCATGATGCTTCTTAATCTAATAGAGCTAC 297
Db 181 TTAAATGGTATTAAATTTAGGAGGTATGATGCTTCTTAATCTAATAGAGCTAT 240
QY 298 TGA 300
Db 241 TGA 243

RESULT 4
US-10-322-281-712
; Sequence 712, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 712
; LENGTH: 31926
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-322-281-712

Query Match 53.6%; Score 160.8; DB 19; Length 31926;
Best Local Similarity 86.8%; Pred. No. 3.7e-38;
Matches 177; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 AAAATATATAAAATTAGCCAGCGGTGATGTGTCATGTCCTGTAGTCCACGCTACTCGGGAGG 60
```

```
Db 17249 AAAATACAAAAATTAGCAGCGTGTGGCATCGCTGTAGTCCCGAGTACTCGGGAGG 17308
QY 61 CTGAGGACGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGATAGTGC 120
Db 17309 CTGAGCAGAGAAATCGCTTGAACCCAGAGATGGAGGTTGCAGTGCAGCGAGATCGTGC 17368
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGCAGACTCTGTCTCAAAAAATAAAAAATAAATA 180
Db 17369 CACTGCACCTCCAGCCTGGGTGACAGAGTGTGACTCTGTCTCAAAAAATAAAAAATAAATA 17428
QY 181 AAAATGCGACGTGATTCAGCA 204
Db 17429 GAAATGGCTAAGTGTGAGGCAGAA 17452
```

```
RESULT 5
US-10-027-632-128663/c
; Sequence 128663, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128663
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128663
```

```
Query Match 52.2%; Score 156.6; DB 13; Length 552;
Best Local Similarity 89.8%; Pred. No. 1.2e-37;
Matches 168; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGCGGTGATGTCATGCTGTAGTCCCGAGTACTCGGGAGG 60
Db 378 AAAATACAAAAATTAGCAGCGATGTGGCATGTGCTGTAGTCCCGAGTACTCGGGAGG 319
QY 61 CTGAGGACGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGATAGTGC 120
Db 318 CTGAGGACGAGAAATCCTCGAACCCAGGAGCGGCTGCAGTGCAGCAAGATGTTGC 259
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGCAGACTCTGTCTCAAAAAATAAAAAATAAATA 180
Db 258 CACTGCACCTCCAGCCTGGGTGACAGAGTGTGACTCTGTCTCAAAAAATAAAAAATAAATA 199
QY 181 AAAATG 187
Db 198 GATATTG 192
```

```
RESULT 6
US-10-027-632-128663/c
; Sequence 128663, Application US/10027632
```

```
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128663
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-128663
```

```
Query Match 52.2%; Score 156.6; DB 17; Length 552;
Best Local Similarity 89.8%; Pred. No. 1.2e-37;
Matches 168; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 AAAATATAAAATAGCCAGCGGTGATGTCATGCTGTAGTCCCGAGTACTCGGGAGG 60
Db 378 AAAATACAAAAATTAGCAGCGATGTGGCATGTGCTGTAGTCCCGAGTACTCGGGAGG 319
QY 61 CTGAGGACGAGAACCTTTGAATCCAGGAGCGCAGGTTGCAGTGCAGCAGATAGTGC 120
Db 318 CTGAGGACGAGAAATCCTCGAACCCAGGAGCGGCTGCAGTGCAGCAAGATGTTGC 259
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGCAGACTCTGTCTCAAAAAATAAAAAATAAATA 180
Db 258 CACTGCACCTCCAGCCTGGGTGACAGAGTGTGACTCTGTCTCAAAAAATAAAAAATAAATA 199
QY 181 AAAATG 187
Db 198 GATATTG 192
```

```
RESULT 7
US-09-764-847-1214/c
; Sequence 1214, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1214
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1214
```

```
Query Match 52.1%; Score 156.4; DB 9; Length 5304;
Best Local Similarity 81.5%; Pred. No. 3.8e-37;
Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
```

Tue Aug 9 17:01:28 2005

```
QY 1 AAAATATATAAAATTAGCCAGGCGTGATGTCATGTGCGCTGTGTAGTCCAGCTACTCTCGGAGG 60
Db 3630 AAACATATAAAATTAGCTGGGCGTGGTGCGACGTGCTGTAGTCCAGCTACTCTCGGAGG 3571
QY 61 CTGAGGCGAGGAACTCTTGAATCCAGGAGCGCAGGTTCAGTGAGCAGAGATAGTGC 120
Db 3570 CTGAGGCGAGGAAATCGCTTAAACCCGGAGCGGAGATTGAGTGAGCTGAGATCATGC 3511
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
Db 3510 CACTGCACCTCCAGCCTGGSCGACAGATGAGGCTCTGCTCAAAAAATAAAATA 3451
QY 181 AAAATGCGAGCTGTGATTACAGAGGTCTGGGTTGAAGCCCA 222
Db 3450 AAAGAAGAAGAAGAAGAGGCTGTCTCTGGTCAAGGCCA 3409
RESULT 8
US-09-764-847-1215/c
; Sequence 1215, Application US/09764847
; Patent No. US20020132767A1
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1215
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-847-1215
Query Match 52.1%; Score 156.4; DB 9; Length 5304;
Best Local Similarity 81.5%; Pred. No. 3.8e-37;
Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 AAAATATATAAAATTAGCCAGGCGTGATGTCATGTGCGCTGTGTAGTCCAGCTACTCTCGGAGG 60
Db 3630 AAACATATAAAATTAGCTGGGCGTGGTGCGACGTGCTGTAGTCCAGCTACTCTCGGAGG 3571
QY 61 CTGAGGCGAGGAACTCTTGAATCCAGGAGCGCAGGTTCAGTGAGCAGAGATAGTGC 120
Db 3570 CTGAGGCGAGGAAATCGCTTAAACCCGGAGCGGAGATTGAGTGAGCTGAGATCATGC 3511
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
Db 3510 CACTGCACCTCCAGCCTGGSCGACAGATGAGGCTCTGCTCAAAAAATAAAATA 3451
QY 181 AAAATGCGAGCTGTGATTACAGAGGTCTGGGTTGAAGCCCA 222
Db 3450 AAAGAAGAAGAAGAAGAGGCTGTCTCTGGTCAAGGCCA 3409
RESULT 9
US-10-092-154-1214/c
; Sequence 1214, Application US/10092154
; Patent No. US20030054375A1
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1214
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-092-154-1214
Query Match 52.1%; Score 156.4; DB 14; Length 5304;
Best Local Similarity 81.5%; Pred. No. 3.8e-37;
Matches 181; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 AAAATATATAAAATTAGCCAGGCGTGATGTCATGTGCGCTGTGTAGTCCAGCTACTCTCGGAGG 60
Db 3630 AAACATATAAAATTAGCTGGGCGTGGTGCGACGTGCTGTAGTCCAGCTACTCTCGGAGG 3571
QY 61 CTGAGGCGAGGAACTCTTGAATCCAGGAGCGCAGGTTCAGTGAGCAGAGATAGTGC 120
Db 3570 CTGAGGCGAGGAAATCGCTTAAACCCGGAGCGGAGATTGAGTGAGCTGAGATCATGC 3511
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATA 180
Db 3510 CACTGCACCTCCAGCCTGGSCGACAGATGAGGCTCTGCTCAAAAAATAAAATA 3451
QY 181 AAAATGCGAGCTGTGATTACAGAGGTCTGGGTTGAAGCCCA 222
Db 3450 AAAGAAGAAGAAGAAGAGGCTGTCTCTGGTCAAGGCCA 3409
RESULT 11
US-10-027-632-128662/c
; Sequence 128662, Application US/10027632
; Patent No. US20020198371A1
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1214
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-092-154-1214/c
; Sequence 1215, Application US/10092154
; Patent No. US20030054375A1
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1214
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-092-154-1215/c
; Sequence 1215, Application US/10092154
; Patent No. US20030054375A1
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1215
; LENGTH: 5304
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-092-154-1215
```

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128662
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-128662
```

```
Query Match 51.9%; Score 155.8; DB 13; Length 556;
Best Local Similarity 90.7%; Pred. No. 2.2e-37;
Matches 166; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60
Db 376 AAAATACAAAAATTAGCCAGGCGATGTTGGCATGTGCTGTAGTCCAGCTACTCGGGAGG 317
QY 61 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 120
Db 316 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGGAGGCTGCAGTGAGCCAGATGGTGC 257
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 256 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 197
QY 181 AAA 183
Db 196 ATA 194
```

```
RESULT 12
US-10-027-632-128662/c
; Sequence 128662, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 128662
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-128662

Query Match 51.9%; Score 155.8; DB 17; Length 556;
Best Local Similarity 90.7%; Pred. No. 2.2e-37;
Matches 166; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60
Db 376 AAAATACAAAAATTAGCCAGGCGATGTTGGCATGTGCTGTAGTCCAGCTACTCGGGAGG 317
QY 61 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 120
Db 316 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGGAGGCTGCAGTGAGCCAGATGGTGC 257
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 256 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 197
QY 181 AAA 183
Db 196 ATA 194
```

```
RESULT 13
US-10-723-860-3426/c
; Sequence 3426, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIORITY FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3426
; LENGTH: 195917
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-723-860-3426
```

```
Query Match 51.9%; Score 155.8; DB 20; Length 195917;
Best Local Similarity 90.7%; Pred. No. 2.7e-36;
Matches 166; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGTGATGTCATGTGCTGTAGTCCAGCTACTCGGGAGG 60
Db 194067 AAAATACAAAAATTAGCCAGGCGATGTTGGCATGTGCTGTAGTCCAGCTACTCGGGAGG 194008
QY 61 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGCAGGTTGCAGTGACGAGATAGTGC 120
Db 194007 CTGAGGCAGGAGCACTCTTGAATCCAGGAGCGGAGGCTGCAGTGAGCCAGATGGTGC 193948
QY 121 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 180
Db 193947 CACTGCACCTCCAGCCTGGGTGACAGAGTGAGACTCTGTCTCAAAAAATAAAATAAATA 193888
QY 181 AAA 183
Db 193887 ATA 193885
```

```
RESULT 14
US-10-027-632-128664/c
```

Tue Aug 9 17:01:28 2005

us-09-899-276c-6.rnpb

```
; Sequence 128664, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128664
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-128664

Query Match      51.8%; Score 155.4; DB 13; Length 556;
Best Local Similarity 90.2%; Pred. No. 2.9e-37;
Matches 165; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCTCGGAGG 60
   |||||
Db 376 AAAATACAAAAATTAGCCAGGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCTCGGAGG 317
   |||||

QY 61 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGGAGTTGCAGTGACGAGATAGTGC 120
   |||||
Db 316 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGGAGTTGCAGTGACGAGATAGTGC 257
   |||||

QY 121 CACTGCACTCCAGCTCGGCTGACAGAGTGAGACTCTCTCTCAAAAAAATAAAAAATA 180
   |||||
Db 256 CACTGCACTCCAGCTCGGCTGACAGAGTGAGACTCTCTCTCAAAAAAATAAAAAATA 197
   |||||

QY 181 AAA 183
   |||
Db 196 ATA 194

Search completed: August 4, 2005, 15:54:26
Job time : 539.142 secs

; Sequence 128664, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128664
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-128664

Query Match      51.8%; Score 155.4; DB 13; Length 556;
Best Local Similarity 90.2%; Pred. No. 2.9e-37;
Matches 165; Conservative 1; Mismatches 17; Indels 0; Gaps 0;

QY 1 AAAATATAAAATTAGCCAGGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCTCGGAGG 60
   |||||
Db 376 AAAATACAAAAATTAGCCAGGCGGTGATGTCATGTGCTGTAGTCCAGCTACTCTCGGAGG 317
   |||||

QY 61 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGGAGTTGCAGTGACGAGATAGTGC 120
   |||||
Db 316 CTGAGGCGAGGAGAACCTTTGAATCCAGGAGGCGGAGTTGCAGTGACGAGATAGTGC 257
   |||||

QY 121 CACTGCACTCCAGCTCGGCTGACAGAGTGAGACTCTCTCTCAAAAAAATAAAAAATA 180
   |||||
Db 256 CACTGCACTCCAGCTCGGCTGACAGAGTGAGACTCTCTCTCAAAAAAATAAAAAATA 197
   |||||

QY 181 AAA 183
   |||
Db 196 ATA 194

RESULT 15
US-10-027-632-128664/c
; Sequence 128664, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 3, 2005, 21:53:02 ; Search time 2612.63 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-7  
Perfect score: 650  
Sequence: 1 ggctgtgcgagatgtcc.....aattgaatgcgtccaccaa 650

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_cm.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	650	100.0	650	6	AX343334 Sequence
2	650	100.0	3227	9	D26087 Human gene
3	616.6	94.9	11793	9	Y18933 Homo sapien
4	615	94.6	9174	9	AF519531 Homo sapi
5	615	94.6	147416	9	AC005549 Homo sapi
6	608.6	93.6	3221	9	AY357296 Homo sapi
7	425.2	65.4	73806	2	AC021520 Homo sapi
8	401.8	61.8	5926	6	AX346555 Sequence
9	365.4	56.2	5926	6	AX346554 Sequence
10	147.8	22.7	316	9	AF493701 Callithri
11	141.4	21.8	307	9	AF493698 Homo sapi
12	139	21.4	312	9	AF493700 Macaca ra
13	138.2	21.3	310	9	AF493697 Homo sapi
14	126.4	19.4	312	9	AF493699 Papio ham
15	94.4	14.5	2788	10	U12470 Mus musculu
16	94.4	14.5	180944	2	AC073824 Mus muscu
17	94.4	14.5	222121	10	AC012294 Mus muscu
18	94.4	14.5	223726	10	AL626807 Mouse DNA
19	94.4	14.5	240425	10	CNS07YOT Mus muscu

20	91.4	14.1	242335	2	AC123203	AC123203 Rattus no
21	91.4	14.1	245808	2	AC114440	AC114440 Rattus no
22	87.4	13.4	28613	10	MMU238892	AJ238892 Mus muscu
23	82.4	12.7	160869	2	AC022399	AC022399 Mus muscu
24	80.4	12.4	3857	10	AF079313	AF079313 Rattus no
25	43.2	6.6	2000	6	AX655393	AX655393 Sequence
26	42.8	6.6	271586	2	AC126739	AC126739 Rattus no
27	42.4	6.5	227406	2	AC118439	AC118439 Rattus no
28	42.4	6.5	301536	2	AC134158	AC134158 Rattus no
29	38.8	6.0	149109	2	AC101018	AC101018 Rattus no
30	38.8	6.0	193472	2	AC115173	AC115173 Rattus no
31	38.8	6.0	240853	2	AC130030	AC130030 Rattus no
32	38.6	5.9	958	1	AF269026	AF269026 Unculture
33	38.6	5.9	214824	2	AC106466	AC106466 Rattus no
34	38.6	5.9	220290	2	BX936393	BX936393 Danio rer
35	38.6	5.9	286016	2	AC112101	AC112101 Rattus no
36	38.6	5.9	323184	2	AC103450	AC103450 Rattus no
37	38.4	5.9	125020	9	AF429315	AF429315 Homo sapi
38	38.4	5.9	176480	2	AC130011	AC130011 Rattus no
39	38.4	5.9	192238	2	AC119537	AC119537 Rattus no
40	38.4	5.9	342774	2	AC116082	AC116082 Rattus no
41	38.2	5.9	6557	6	AX345902	AX345902 Sequence
42	38.2	5.9	64990	9	AP000251	AP000251 Homo sapi
43	38.2	5.9	100000	9	AP000030	AP000030 Homo sapi
44	38.2	5.9	100000	9	AP000212	AP000212 Homo sapi
45	38.2	5.9	223233	2	AC097562	AC097562 Rattus no

ALIGNMENTS

RESULT 1  
AX343334  
LOCUS AX343334  
DEFINITION Sequence 7 from Patent EP1170372.  
ACCESSION AX343334  
VERSION AX343334.1  
KEYWORDS GI:18491684  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Roesl,F., Soto,U., Coy,J., Finzer,P., Delius,H., Poustka,A., zur Hausen,H. and Patzelt,A.  
TITLE Regulatory sequences of the human mcp-1 gene  
JOURNAL Patent: EP 1170372-A 7 09-JAN-2002;  
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS (DE)

FEATURES  
Location/Qualifiers  
1..650  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN	Query Match	100.0%	Score 650;	DB 6;	Length 650;
Best Local Similarity	100.0%	Pred. No. 4.6e-190;	Mismatches	0;	Indels
Matches	650;	Conservative	0;	Gaps	0;
QY	1	GGCTTTGTGCGGAGATGTTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCGGGCCCCA	60		
Db	1	GGCTTTGTGCGGAGATGTTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCGGGCCCCA	60		
QY	61	GTATCTGGATCAGGCTCCAGCCAAATGTCATCTTCTACGGATCTGGGAACCTTCCA	120		
Db	61	GTATCTGGATCAGGCTCCAGCCAAATGTCATCTTCTACGGATCTGGGAACCTTCCA	120		
QY	121	AAGCTGCTCTCCAGAGTGGGAATTTCCACTCACTTCTCTACGGCAGCACTGACCTCCC	180		
Db	121	AAGCTGCTCTCCAGAGTGGGAATTTCCACTCACTTCTCTACGGCAGCACTGACCTCCC	180		
QY	181	AGCGGGGGAGGGGCATCTTTTCTTGACAGACAGAAAGTGGGAGGACAGACAGCTGTCACTTT	240		

191	AGCGGGGAGGGCATCTTTCTTGACAGACAGAACTGGAGGCAGACAGCTGTCACTTT	240
241	CCAGAAGACTTTCTTTTCTGATTCATACCTTCACCTTCCTGTGTTTACTGCTGATAT	300
241	CCAGAAGACTTTCTTTTCTGATTCATACCTTCACCTTCCTGTGTTTACTGCTGATAT	300
301	ATGCNAAAGCCAAAGTCACTTTTCAGAGATCACAACTCCTTCCTGAAGTAGAGACATGCTT	360
301	ATGCNAAAGCCAAAGTCACTTTTCAGAGATGACAACTCCTTCCTGAAGTAGAGACATGCTT	360
361	CCAAACACTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGAAAGTTTTCTCTGTGA	420
361	CCAAACACTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGAAAGTTTTCTCTGTGA	420
421	CCATGGGCTAATTGGTCTCCTTCTCTGGATTGTGGCTTATCAGATAAAACAAGTGAGTC	480
421	CCATGGGCTAATTGGTCTCCTTCTCTGGATTGTGGCTTATCAGATAAAACAAGTGAGTC	480
481	ATGCCACAGGATGCTATAAAGCCCAATTGATTTCTGGGATTCATAGTAGTGATGCTGATATGA	540
481	ATGCCACAGGATGCTATAAAGCCCAATTGATTTCTGGGATTCATAGTAGTGATGCTGATATGA	540
541	CTAAGCCAGGAGAGACTTAATTTAAAGATCTCAGCATCTTTCAAGCTTTGTTAACTAGAGAA	600
541	CTAAGCCAGGAGAGACTTAATTTAAAGATCTCAGCATCTTTCAAGCTTTGTTAACTAGAGAA	600
601	AAACCCGAAGCATGACTGGATTATTAAGGGAAATTTGAATGCGGTCCACCA	650
601	AAACCCGAAGCATGACTGGATTATTAAGGGAAATTTGAATGCGGTCCACCA	650

RESULT 2	
HUMMCP1	
LOCUS	3227 bp DNA linear PRI 14-APR-2000
DEFINITION	Human gene for monocyte chemoattractant protein-1, 5'-genome region.
ACCESSION	D26087
VERSION	D26087.1 GI:516772
KEYWORDS	MCP-1; monocyte chemoattractant protein-1.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. 1 (sites) Rollins,B.J., Stier,P., Ernst,T. and Wong,G.G. The human homolog of the JE gene encodes a monocyte secretory protein JOURNAL Mol. Cell. Biol. 9 (11), 4687-4695 (1989)
JOURNAL MEDLINE	90097880
PUBMED	2513477
REFERENCE	2 (sites) Shyy,Y.J., Li,Y.S. and Kolattukudy,P.E. Structure of human monocyte chemotactic protein gene and its regulation by TPA JOURNAL Biochem. Biophys. Res. Commun. 169 (2), 346-351 (1990)
JOURNAL MEDLINE	90290466
PUBMED	2357211
REFERENCE	3 (bases 1 to 3227) Ueda,A., Okuda,K., Ohno,S., Shirai,A., Igarashi,T., Matsunaga,K., Fukushima,J., Kawamoto,S., Ishigatsubo,Y. and Okubo,T. NF-kappa B and Sp1 regulate transcription of the human monocyte chemoattractant protein-1 gene J. Immunol. 153 (5), 2052-2063 (1994)
TITLE	
JOURNAL MEDLINE	94327939
PUBMED	8051410
REFERENCE	4 (bases 1 to 3227) Ueda,A. Direct Submission Submitted (06-DEC-1993) Atsuhisa Ueda, Yokohama City University School of Medicine, First Department of Internal Medicine; 3-9 Fukuura, Kanazawa-ku, Yokohama 236, Japan (Tel:045-787-2630, Fax:045-786-3444)
AUTHORS	
TITLE	
JOURNAL	

```

COMMENT      Submitted (06-Dec-1993) to DDBJ by:
              Atsuhisa Ueda
              First Department of Internal Medicine
              Yokohama City University School of Medicine
              3-9 Fukuura, Kanazawa-ku
              Yokohama 236
              Japan
              Phone: 045-787-2630
              Fax: 045-786-3444.
FEATURES     Location/Qualifiers
              1..3227
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /tissue_type="placenta"
                /clone_lib="EMBL3 SP6/T7"
                400..407
                  /note="kappa B"
              enhancer
                2948..2953
                GC_signal
                TATA_signal
                exon
                  2975..2979
                  3012..3147
                  /number=1
                  3072..3147
                  /codon_start=1
                  /product="monocyte chemoattractant protein-1"
                  /protein_id="BA005080.1"
                  /db_xref="GI:992947"
                  /translation="MKVSAALLCLLLIAATFIPQGLAQP"
                  3148..3227
                  /number=1
              intron

```

Query Match	100.0%;	Score 650;	DB 9;	Length 3227;
Best Local Similarity	100.0%;	Pred. No. 5e-190;		
Matches 650;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GGCTTGCCGAGATGTTCCACGACAGCCCATGTGAGAGCTCCCTGGCTCGCGGCCCA	60	
DB	262	GGCCTTGCCGAGATGTTCCACGACAGCCCATGTGAGAGCTCCCTGGCTCGCGGCCCA	321	
QY	61	GTATCTGGAATGCAGGCTCCACGCCAAATGCATTCTTTACGGGATCTGGGAACCTTCCA	120	
DB	322	GTATCTGGAATGCAGGCTCCACGCCAAATGCATTCTTTACGGGATCTGGGAACCTTCCA	381	
QY	121	AAGCTGCCTCCTCAGATGGGAATTTCCACTCACTTCTCTCACGCGACGACTGCACCTCCC	180	
DB	382	AAGCTGCCTCCTCAGATGGGAATTTCCACTCACTTCTCTCACGCGACGACTGCACCTCCC	441	
QY	181	ACGGGGGAGGGCATCTTTTCTTGACAGACAGAAGTGGGAGGACAGACAGCTGTCACTTT	240	
DB	442	ACGGGGGAGGGCATCTTTTCTTGACAGACAGAAGTGGGAGGACAGACAGCTGTCACTTT	501	
QY	241	CCAGAAGACTTTCTTTTCTGATTCATAACCTTCACTTCCCTGTGTGTTTACTGTCTGATAT	300	
DB	502	CCAGAAGACTTTCTTTTCTGATTCATAACCTTCACTTCCCTGTGTGTTTACTGTCTGATAT	561	
QY	301	ATGCAAAGCCCAAGTCACTTTCCAGAGATGACAACTCTCTCTGAAGTAGAGACATGCTT	360	
DB	562	ATGCAAAGCCCAAGTCACTTTCCAGAGATGACAACTCTCTCTGAAGTAGAGACATGCTT	621	
QY	361	CAACAACCTCAGAAGCCTATGTGAACACTCAGCCAGCAAGCTGGGAAGTTTTTCTCTGTGA	420	
DB	622	CAACAACCTCAGAAGCCTATGTGAACACTCAGCCAGCAAGCTGGGAAGTTTTTCTCTGTGA	681	
QY	421	CCATGGGCTAATTGGTCTCCTTCTCTGGATTCGGCTTATCAGATAAAAACAAGTGAATC	480	
DB	682	CCATGGGCTAATTGGTCTCCTTCTCTGGATTCGGCTTATCAGATAAAAACAAGTGAATC	741	
QY	481	ATGCCACAGGATGCTCTATAAGGCCATTGATTCTTGGGATTTCTATGAGTGATGCTGATAGA	540	
DB	742	ATGCCACAGGATGCTCTATAAGGCCATTGATTCTTGGGATTTCTATGAGTGATGCTGATAGA	801	
QY	541	CTAAGCCAGGAGAGACTTATTTAAAGATCTCAGCATCTTTCACTTGTTTAACTAGAGAA	600	

Db	802	CTAAGCCAGGAGACTTATTTAAAGATCTCAGCATCTTTCAGCTTGTTAACTAGAGAA	861
QY	601	AAACCCAGCATGACTGGATTATAAGGGAATTAAGTGGTCCACCAA	650
Db	862	AAACCCAGCATGACTGGATTATAAGGGAATTAAGTGGTCCACCAA	911
RESULT 3			
HSY18933			
LOCUS	HSY18933	11793 bp	DNA linear PRI 18-OCT-2000
DEFINITION	Homo sapiens MCP-1 gene and enhancer region.		
ACCESSION	Y18933		
VERSION	Y18933.1	GI:10933860	
KEYWORDS	MCP-1 gene; monocyte chemoattractant protein-1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Finzer, P., Soto, U., Dellus, H., Patzelt, A., Coy, J. F., Poustka, A., zur Hausen, H. and Roel, F.		
TITLE	Differential transcriptional regulation of the monocyte-chemoattractant protein-1 (MCP-1) gene in tumorigenic and non-tumorigenic HPV 18 positive cells: the role of the chromatin structure and AP-1 composition		
JOURNAL	Oncogene 19 (29), 3235-3244 (2000)		
MEDLINE	20374005		
PUBMED	10918580		
REFERENCE	2	(bases 1 to 11793)	
AUTHORS	Roel, F.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-APR-1999) F. Roel, Applied Tumor Virology, DKPZ (German Cancer Research Institute), Im Neuenheimer Feld 506, 69120 Heidelberg, FRG		
FEATURES			
source	1. .11793		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
gene	4906. .9073		
	/gene="MCP-1"		
enhancer	4906. .4915		
	/gene="MCP-1"		
enhancer	4935. .4944		
	/gene="MCP-1"		
TATA_signal	7513. .7517		
	/gene="MCP-1"		
mRNA	join(<7550. .7685,8468. .8585,8968. .>9073)		
exon	<7550. .7685		
	/gene="MCP-1"		
CDS	/number=1		
	join(7610. .7685,8468. .8585,8968. .9073)		
	/gene="MCP-1"		
	/codon_start=1		
	/product="monocyte chemoattractant protein-1"		
	/protein_id="CAG14049.1"		
	/db_xref="GI:10933861"		
	/db_xref="GOA:P13500"		
	/translation="MKVSAALLCLLLIAATFIPOGLAOPDAINAPVTCYFNTRKIS VORLASRYRITSSKCPKEAVIFKTIKAVEICADPKQKWQDSMDHLDKQTQPKT"		
intron	7686. .8467		
	/gene="MCP-1"		
exon	8468. .8585		
	/number=1		
	/gene="MCP-1"		
intron	8586. .8967		
	/number=2		
	/gene="MCP-1"		

exon	8968. .>9073		
	/gene="MCP-1"		
	/number=3		
ORIGIN			
Query Match	94.9%;	Score 616.6;	DB 9; Length 11793;
Best Local Similarity	97.7%;	Pred. No. 1.2e-179;	
Matches	636;	Conservative	0; Mismatches 14; Indels 1; Gaps 1;
QY	1	GGCTGTGCGGAGATGTTCCAGCAGCAGCCCATGTGAGAGCTCCCTGGCTCGGGGCCCA	60
Db	4797	GGCTGTGCGGAGATGTTCCAGCAGCAGCCCATGTGAGAGCTCCCTGGCTCGGGGCCCA	4856
QY	61	GTATCTGGAATCAGGCTCCAGCCAAATGCATTCTCTTACGGGATCTGGGAACCTTCCA	120
Db	4857	GTATCTGGAATCAGGCTCCAGCCAAATGCATTCTCTTACGGGATCTGGGAACCTTCCA	4916
QY	121	AGCTGCTCTCCTCAGAGTGGGAATTTCCACTCACTCTCTCAGCCAGCAGCTGACCTCCC	180
Db	4917	AAAGCTGCTCTCCTCAGAGTGGGAATTTCCACTCACTCTCTCAGCCAGCAGCTGACCTCCC	4976
QY	181	ACGGGGGAGGGCATCTTTTCTTGACAGACAGAGTGGGAGCAGACAGCTGTCACTTT	240
Db	4977	ACGGGGGAGGGCATCTTTTCTTGACAGACAGAGTGGGAGCAGACAGCTGTCACTTT	5036
QY	241	CCAGAAGACTTTTCTTTCTGATTACACCTTCACCTTCCTCTGTGTGTTACTGTCTGATAT	300
Db	5037	CCAGAAGACTTTTCTTTCTGATTACACCTTCACCTTCCTCTGTGTGTTACTGTCTGATAT	5096
QY	301	ATGCAAGGCCAAAGTCACCTTTCAGAGATGACAACTCTCTCTGAAGTAGAGACATGCTT	360
Db	5097	ATGCAAGGCCAAAGTCACCTTTCAGAGATGACAACTCTCTCTGAAGTAGAGACATGCTT	5156
QY	361	CAACACACTCAGAAGCTATGTGAACACTCAGCAGCAAGCT-GGAAGTTTTTCTCTGTG	419
Db	5157	CAACACACTCAGAAGCTATGTGAACACTCAGCAGCAAGCTGGAAGTTTTTCTCTGTG	5216
QY	420	ACCATGGGCTAAATGGTCTCTCTTCTTGATTGTGGCTTATCAGATAAAACAAGTGAGT	479
Db	5217	ACCATGGGCTAAATGGTCTCTCTTCTTGATTGTGGCTTATCAGATAAAACAAGTGAGT	5276
QY	480	CATGCCACAGGATGTCATTAAGCCCATTAATCTGGGATCTATGAGTGATGCTGATATG	539
Db	5277	CATGCCACAGGATGTCATTAAGCCCATTAATCTGGGATCTATGAGTGATGCTGATATG	5336
QY	540	ACTAAGCCAGGAGAGACTTATTTAAGATCTCAGACATCTTTCAGCTTGTTAACCTAGAGA	599
Db	5337	ACTAAGCCAGGAGAGACTTATTTAAGATCTCAGACATCTTTCAGCTTGTTAACCTAGAGA	5396
QY	600	AAACCCGAGCATGACTGGATTATAAGGGAATTAAGTGGTCCACCAA	650
Db	5397	AAACCCGAGCATGACTGGATTATAAGGGAATTAAGTGGTCCACCAA	5447
RESULT 4			
AF519531			
LOCUS	AF519531	9174 bp	DNA linear PRI 09-JUL-2002
DEFINITION	Homo sapiens small inducible cytokine A2 (monocyte chemotactic protein 1) (SCYA2) gene, complete cds.		
ACCESSION	AF519531		
VERSION	AF519531.1	GI:21435976	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	(bases 1 to 9174)	
AUTHORS	Rieder, M. J., Armel, T. Z., Carrington, D. P., Ozuna, M., Kuldane, S. A., Rajkumar, N., Toth, E. J., Yi, Q. and Nickerson, D. A.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUN-2002) Genome Sciences, University of Washington, 1705 NE Pacific, Seattle, WA 98195, USA		
COMMENT	To cite this work please use: SeattleSNPs. NHLBI HL6682 Program		

Tue Aug 9 17:01:28 2005

us-09-899-276c-7.rge

for Genomic Applications, UW-FHCRC, Seattle, WA (URL:

http://pga.gs.washington.edu).

Location/Qualifiers

1. .9174

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

variation

/frequency="0.17"

/replace="g"

variation

/frequency="0.03"

/replace="g"

variation

/frequency="0.15"

/replace="t"

variation

/frequency="0.01"

/replace="g"

variation

/frequency="0.16"

/replace="t"

variation

/frequency="0.02"

/replace="a"

repeat\_region

614. .720

/rpt\_family="MIR"

/rpt\_type=dispersed

variation

/frequency="0.14"

/replace="g"

variation

/frequency="0.01"

/replace="a"

variation

/frequency="0.01"

/replace="t"

variation

/frequency="0.03"

/replace="t"

variation

/frequency="0.23"

/replace="a"

variation

/frequency="0.15"

/replace="g"

variation

/frequency="0.08"

/replace="a"

variation

/frequency="0.04"

/replace="g"

variation

/frequency="0.02"

/replace="c"

repeat\_region

1718. .1830

/rpt\_family="MIR"

/rpt\_type=dispersed

variation

/frequency="0.01"

/replace="c"

variation

/frequency="0.13"

/replace="t"

variation

/frequency="0.02"

/replace="g"

variation

/frequency="0.04"

/replace="g"

variation

/frequency="0.01"

/replace="a"

variation

/frequency="0.08"

/replace="t"

variation

/frequency="0.03"

/replace="a"

variation

/frequency="0.14"

/replace="c"

repeat\_region

3228. .3445

/rpt\_family="MIR"

/rpt\_type=dispersed

variation

/frequency="0.01"

/replace="a"

variation

/frequency="0.25"

/replace="c"

gene

4065. .5981

/gene="SCYA2"

join(4065. .4193,4990. .5107,5490. .5981)

/gene="SCYA2"

/product="small inducible cytokine A2 (monocyte

chemotactic protein 1)"

CDS

join(4118. .4193,4990. .5107,5490. .5595)

/gene="SCYA2"

/codon\_start=1

/product="small inducible cytokine A2 (monocyte

chemotactic protein 1)"

/protein\_id="AAM54046.1"

/db\_xref="GI:21435977"

/translation="MKVSAALCLLLIAATFIQGLAQPDAINAPVTCCYNFTNRKIS

VQRLASVRRITSSKCPKEAVIFKTIIVAKEICADPKQKQVQDSMDHLDKQTPTKT"

4296. .4409

/rpt\_family="MIR"

/rpt\_type=dispersed

misc\_feature

4359. .4462

/gene="SCYA2"

/note="Region not scanned for variation"

variation

4748

/gene="SCYA2"

/frequency="0.26"

/replace="a"

4777. .4814

/rpt\_family="MIR"

/rpt\_type=dispersed

variation

4881

/gene="SCYA2"

/frequency="0.14"

/replace="g"

variation

5018

/gene="SCYA2"

/frequency="0.47"

/replace="c"

variation

5132

/gene="SCYA2"

/frequency="0.01"

/replace="t"

variation

5660

/gene="SCYA2"

/frequency="0.18"

/replace="t"

variation

6184

/frequency="0.01"

/replace="c"

variation

6518

/frequency="0.01"

/replace="t"

variation

6531

/frequency="0.05"

/replace="t"

6837. .6916

/rpt\_family="L2"

variation	/rpt_type=dispersed 7436 /frequency="0.07" /replace="t"	variation	480	CATGCCACAGGATGCTCTATAAGCCCATTTGATTCTGGGATTCCTATGATGCTGATG 539 
variation	7605 /frequency="0.03" /replace="a"	variation	1785	CATGCCACAGGATGCTCTATAAGCCCATTTGATTCTGGGATTCCTATGATGCTGATG 1844 
variation	7614 /frequency="0.01" /replace="a"	QY	540	ACTAAGCCAGGAGAGACTTTATTTAAAGATCTCAGCATCTTTTTCAGCTTTGTTAACTAGAGA 599 
variation	7659 /frequency="0.03" /replace="a"	Db	1845	ACTAAGCCAGGAGAGACTTTATTTAAAGATCTCAGCATCTTTTTCAGCTTTGTTAACTAGAGA 1904 
variation	7710 /frequency="0.05" /replace="a"	QY	600	AAACCCGAGCATGACTGATTTATAAGGGAATTGAATGCCGTCACCA 650 
variation	7843 /frequency="0.31" /replace="c"	Db	1905	AAACCCGAGCATGACTGATTTATAAGGGAATTGAATGCCGTCACCA 1955 
variation	8072 /frequency="0.01" /replace="t"	RESULT 5		
repeat_region	8118..8271 /rpt_family="MaLR" /rpt_type=dispersed	AC005549/c		
repeat_region	8272..8734 /rpt_family="MaLR" /rpt_type=dispersed	LOCUS		
variation	8597 /frequency="0.07" /replace="c"	DEFINITION	147416 bp DNA linear PRI 22-SEP-1998	
repeat_region	8735..9033 /rpt_family="Alu" /rpt_type=dispersed	AC005549	Homo sapiens chromosome 17, clone hRPK.215_E13, complete sequence.	
variation	8775 /frequency="0.07" /replace="a"	AC005549.1	GI:3598724	
		HTG.		
		ORGANISM	Homo sapiens (human)	
		REFERENCE	1 (bases 1 to 147416)	
		AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.	
		TITLE	Homo sapiens chromosome 17, clone hRPK.215_E13	
		JOURNAL	Unpublished	
		REFERENCE	2 (bases 1 to 147416)	
		AUTHORS	Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C., Jaccotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	
		TITLE	Direct Submission	
		JOURNAL	Submitted (27-AUG-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
		REFERENCE	3 (bases 1 to 147416)	
		AUTHORS	Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C., Jaccotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	
		TITLE	Direct Submission	
		JOURNAL	Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
		REFERENCE	4 (bases 1 to 147416)	
		AUTHORS	Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C., Jaccotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	
		TITLE	Direct Submission	
		JOURNAL	Submitted (12-SEP-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
		REFERENCE	4 (bases 1 to 147416)	
		AUTHORS	Birren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K., Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Geraghty,K., Grant,G., Hagos,B., Heaford,A., Herena,L., Horton,L., Howland,J.C., Jaccotot,L., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D., Ye,W.J., Zhao,J. and Zody,M.	

Query Match	94.6%;	Score 615;	DB 9;	Length 9174;
Best Local Similarity	97.5%;	Pred. No. 3.9e-179;		
Matches 635;	Conservative 0;	Mismatches 15;	Indels 1;	Gaps 1;
1	GGCTTGTGCGGAGATGTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCGGGGCCCA	60		
1305	GGCTTGTGCGGAGATGTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCGGGGCCCA	1364		
61	GTATCTGGAATGAGGCTCAGGCCAAATGCATTTCTTCTACGGGATCTGGGAACCTTCCA	120		
1365	GTATCTGGAATGAGGCTCAGGCCAAATGCATTTCTTCTACGGGATCTGGGAACCTTCCA	1424		
121	AAGCTGCCCTCCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCGCAGCACTGACCTCCC	180		
1425	AAGCTGCCCTCCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCGCAGCACTGACCTCCC	1484		
181	AGCGGGGAGGGGATCTTTTCTTGACAGACGAGAGTGGGAGGCAGACAGCTGTCACTTT	240		
1485	AGCGGGGAGGGGATCTTTTCTTGACAGACGAGAGTGGGAGGCAGACAGCTATCACTTT	1544		
241	CGAGAAGACTTTTCTTGTGATTCATACCCCTTCACTTCCCTGTGTGTTTACTGTCTGATAT	300		
1545	CGAGAAGACTTTTCTTGTGATTCATACCCCTTCACTTCCCTGTGTGTTTACTGTCTGATAT	1604		
301	ATGCAAAAGCCCAAGTCACATTTCCAGAGATGACAACTCCCTTCTGTAAGTAGAGACATGCTT	360		
1605	ATGCAAAAGCCCAAGTCACATTTCCAGAGATGACAACTCCCTTCTGTAAGTAGAGACATGCTT	1664		
361	CCAACACTCAGAAAGCCTATGTGAACACTCAGCCAGCAAAAGCT-GGAAGTTTTTCTCTGTG	419		
1665	CCAACACTCAGAAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGGAAGTTTTTCTCTGTG	1724		
420	ACCATGGGCTAATTTGGTCTCTCTTCTCTGGAATGTGGCTTATCAGATAAAAAACAAGTGAGT	479		
1725	ACCATGGGCTAATTTGGTCTCTCTTCTCTGGAATGTGGCTTATCAGATAAAAAACAAGTGAGT	1784		

repeat_region	/rpt_family="MLK"	11567.11836
repeat_region	/rpt_family="L12"	12501.12981
repeat_region	/rpt_family="MER9"	13257.13257
repeat_region	/rpt_family="AluJb"	15248.15364
repeat_region	/rpt_family="MIR"	16653.16653
repeat_region	/rpt_family="MIR"	16830.16949
repeat_region	/rpt_family="MERSA"	17203.17380
repeat_region	/rpt_family="LTR16A"	17393.18306
repeat_region	/rpt_family="HERVL"	18315.24307
repeat_region	/rpt_family="L1PA2"	24309.24335
repeat_region	/rpt_family="AT_rich"	24513.25245
repeat_region	/rpt_family="HERVL"	25613.25758
repeat_region	/rpt_family="HERVL"	26135.26228
repeat_region	/rpt_family="LTR16A"	27950.27950
repeat_region	/rpt_family="MER94"	28930.28968
repeat_region	/rpt_family="AT_rich"	29960.30341
repeat_region	/rpt_family="MLT1I"	32364.32581
repeat_region	/rpt_family="MIR"	32601.32902
repeat_region	/rpt_family="Alusx"	33935.33935
repeat_region	/rpt_family="L1PA2"	39362.39414
repeat_region	/rpt_family="A_rich"	40047.40047
repeat_region	/rpt_family="MIR"	40112.40151
repeat_region	/rpt_family="AT_rich"	40442.40442
repeat_region	/rpt_family="AluJo"	40898.40898
repeat_region	/rpt_family="MLT1J"	41293.41469
repeat_region	/rpt_family="L1PA6"	41956.41956
repeat_region	/rpt_family="purine-rich"	42486.42569
repeat_region	/rpt_family="MLT1I"	42975.43098
repeat_region	/rpt_family="MERSA"	43554.43983
repeat_region	/rpt_family="LTR24"	44781.44811
repeat_region	/rpt_family="AT_rich"	46184.46396
repeat_region	/rpt_family="MIR"	47075.47144
repeat_region	/rpt_family="MLT1G"	47266.47300
repeat_region	/rpt_family="(CA)n"	47444.47444
repeat_region	/rpt_family="MUTIG"	48200.48200

```
10109. .10258  
/rpt_family="MER5A"  
complement(11031. .11072)
```

Matches 635; Conservative 0; Mismatches 15; Indels 1; Gaps 1;									
QY	1	GGCTTGTGCGCAGATGTTCCAGCACAGCCCCATGTGAGAGCTCCCTGGCTCGGGCCCCA	60						
Db	64230	GGCTTGTGCGCAGATGTTCCAGCACAGCCCCATGTGAGAGCTCCCTGGCTCGGGCCCCA	64171						
QY	61	GTATCTGGAATCAGGCTCCAGCCAAATGATCTCTTCTACGGGATCTGGGAACCTTCCA	120						
Db	64170	GTATCTGGAATCAGGCTCCAGCCAAATGATCTCTTCTACGGGATCTGGGAACCTTCCA	64111						
QY	121	AAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCC	180						
Db	64110	AAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCC	64051						
QY	181	AGCGGGGGAGGGCATCTTTTCTTGACAGACAGAGTGGGAGGACAGACACTGTCACTTT	240						
Db	64050	AGCGGGGGAGGGCATCTTTTCTTGACAGACAGAGTGGGAGGACAGACACTGTCACTTT	63991						
QY	241	CCAGAGACATTTCTTTCTGATTCATACCTTCACTTCCCTGTGTTTACTGTCTGATAT	300						
Db	63990	CCAGAGACATTTCTTTCTGATTCATACCTTCACTTCCCTGTGTTTACTGTCTGATAT	63931						
QY	301	ATCCAAAGCCCAAGTCACTTTCCAGAGATGACAACTCCTTCCCTGAAGTAGACATGCTT	360						
Db	63930	ATCCAAAGCCCAAGTCACTTTCCAGAGATGACAACTCCTTCCCTGAAGTAGACATGCTT	63871						
QY	361	CCAACTCTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCT-GGAAGTTTTTCTCTGTG	419						
Db	63870	CCAACTCTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGAAGTTTTTCTCTGTG	63811						
QY	420	ACCATGGCTTAATGGTCTCTCTCTGGAATTTGGCTTATCAGATAAAACAAGTGAGT	479						
Db	63810	ACCATGGCTTAATGGTCTCTCTCTGGAATTTGGCTTATCAGATAAAACAAGTGAGT	63751						
QY	480	CATGCCACAGATGCTCTTAAGCCCATTTGATCTGGGATCTATGATGATGCTGATATG	539						
Db	63750	CATGCCACAGATGCTCTTAAGCCCATTTGATCTGGGATCTATGATGATGCTGATATG	63691						
QY	540	ACTAAGCCAGGAGAGACTTATTTAAAGATCTCAGCATCTTTGAGCTTTGTAACCTAGAGA	599						
Db	63690	ACTAAGCCAGGAGAGACTTATTTAAAGATCTCAGCATCTTTGAGCTTTGTAACCTAGAGA	63631						
QY	600	AAACCCGAGCATGACTGGATTATAAGGGAAATTTGAATGCGGTCCACCAA	650						
Db	63630	AAACCCGAGCATGACTGGATTATAAGGGAAATTTGAATGCGGTCCACCAA	63580						
RESULT 6									
AY357296									
LOCUS									
Homo sapiens MCP1 (MCP1) gene, promoter region and partial cds.									
DEFINITION									
Homo sapiens MCP1 (MCP1) gene, promoter region and partial cds.									
ACCESSION									
AY357296									
VERSION									
AY357296.1 GI:34559719									
KEYWORDS									
Homo sapiens (human)									
SOURCE									
Homo sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE									
1 (bases 1 to 3221)									
Nyquist,P.A. and Degraba,T.J.									
Two New Single Nucleotide Polymorphisms (SNPs) in the MCP-1									
Promoter in Patients with Carotid Atherosclerosis: Transcriptional									
Induction and New Protein Binding Sites									
JOURNAL									
Unpublished									
2 (bases 1 to 3221)									
Nyquist,P.A. and Degraba,T.J.									
Direct Submission									
TITLE									
Submitted (31-JUL-2003) Medicine, Inova Fairfax, 3300 Gallows Road,									
Falls Church, VA 22402-3100, USA									
JOURNAL									
Location/Qualifiers									
FEATURES									
1..3221									
source									
/organism="Homo sapiens"									
/mol_type="genomic DNA"									

/db_xref="taxon:9606"									
225_..>3221									
/gene="MCP1"									
/note="HSMCR30; MCP-1"									
225_..2746									
/gene="MCP1"									
225									
/gene="MCP1"									
/note="SNP"									
/replace="c"									
688									
/gene="MCP1"									
/note="SNP"									
/replace="t"									
1878									
/gene="MCP1"									
/note="SNP"									
/replace="c"									
2445									
/gene="MCP1"									
/note="SNP"									
/replace="g"									
2710_..2714									
/gene="MCP1"									
/evidence=not_experimental									
2747_..2882									
/gene="MCP1"									
/product="MCP1"									
/evidence=not_experimental									
2807_..2882									
/gene="MCP1"									
/codon_start=1									
/evidence=not_experimental									
/product="MCP1"									
/protein_id="AAQ75526.1"									
/db_xref="GI:34559720"									
/translation="MKVSAALLCLLLIAATFIPOGLAQP"									
ORIGIN									
Query Match 93.6%; Score 608.6; DB 9; Length 3221;									
Best Local Similarity 97.7%; Pred. No. 3.5e-177;									
Matches 628; Conservative 0; Mismatches 14; Indels 1; Gaps 1;									
QY	9	CCGAGATGTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG	68						
Db	1	CCGAGATGTTCCAGCACAGCCCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG	60						
QY	69	AATCAGGCTCCAGCCAAATGATCTCTTCTACGGGATCTGGGAATCTTCCAAAGCTGCC	128						
Db	61	AATCAGGCTCCAGCCAAATGATCTCTTCTACGGGATCTGGGAATCTTCCAAAGCTGCC	120						
QY	129	TCCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGG	188						
Db	121	TCCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGG	180						
QY	189	AGGGCATCTTTTCTTGACAGACAGAGTGGGAGGACAGACAGCTGTCACTTTCCAGAA	248						
Db	181	AGGGCATCTTTTCTTGACAGACAGAGTGGGAGGACAGACAGCTGTCACTTTCCAGAA	240						
QY	249	CTTTCTTTTCTGATTCATACCTTCACTTCCCTGTGTTTACTGTCTGATATATGCAAA	308						
Db	241	CTTTCTTTTCTGATTCATACCTTCACTTCCCTGTGTTTACTGTCTGATATATGCAAA	300						
QY	309	GCCAGTCACTTTCCAGAGATGCAACTCTTCTCTGAAGTAGAGACATGCTTCCACACT	368						
Db	301	GCCAGTCACTTTCCAGAGATGCAACTCTTCTCTGAAGTAGAGACATGCTTCCACACT	360						
QY	369	CAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCT-GGAAGTTTTTCTCTGTGACCAT	427						
Db	361	CAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGAAGTTTTTCTCTGTGACCAT	420						
QY	428	CTAATTGGTCTCTCTCTCTGGATTGTGGCTTATCAGATAAAACAAGTGAGTCA	487						

- \* will be sequenced to completion. In the event that
- \* the record is updated, the accession number will
- \* be preserved.

*	1	862:	contig of 862 bp in length
*	863	962:	gap of 100 bp
*	963	1850:	contig of 888 bp in length
*	1851	1950:	gap of 100 bp
*	1951	2817:	contig of 867 bp in length
*	2818	2917:	gap of 100 bp
*	2918	3788:	contig of 871 bp in length
*	3789	3888:	gap of 100 bp
*	3889	4760:	contig of 872 bp in length
*	4761	4860:	gap of 100 bp
*	4861	5725:	contig of 865 bp in length
*	5726	5823:	gap of 100 bp
*	5826	6720:	contig of 895 bp in length
*	6721	6820:	gap of 100 bp
*	6821	7680:	contig of 860 bp in length
*	7681	7780:	gap of 100 bp
*	7781	8666:	contig of 886 bp in length
*	8667	8666:	gap of 100 bp
*	8767	9637:	contig of 871 bp in length
*	9638	9737:	gap of 100 bp
*	9738	10614:	contig of 877 bp in length
*	10615	10714:	gap of 100 bp
*	10715	11607:	contig of 893 bp in length
*	11608	11707:	gap of 100 bp
*	11708	12583:	contig of 876 bp in length
*	12584	12683:	gap of 100 bp
*	12684	13570:	contig of 887 bp in length
*	13571	13670:	gap of 100 bp
*	13671	14560:	contig of 890 bp in length
*	14561	14660:	gap of 100 bp
*	14661	15336:	contig of 876 bp in length
*	15337	15636:	gap of 100 bp
*	15637	16521:	contig of 883 bp in length
*	16522	16621:	gap of 100 bp
*	16622	17518:	contig of 897 bp in length
*	17519	17618:	gap of 100 bp
*	17619	18516:	contig of 898 bp in length
*	18517	18616:	gap of 100 bp
*	18617	19500:	contig of 884 bp in length
*	19501	19600:	gap of 100 bp
*	19601	20455:	contig of 855 bp in length
*	20456	20555:	gap of 100 bp
*	20556	21441:	contig of 886 bp in length
*	21442	21541:	gap of 100 bp
*	21542	22418:	contig of 877 bp in length
*	22419	22518:	gap of 100 bp
*	22519	23376:	contig of 858 bp in length
*	23377	24376:	gap of 100 bp
*	24377	24342:	contig of 866 bp in length
*	24343	24442:	gap of 100 bp
*	24443	25233:	contig of 791 bp in length
*	25234	25333:	gap of 100 bp
*	25334	26223:	contig of 890 bp in length
*	26224	26323:	gap of 100 bp
*	26324	27194:	contig of 871 bp in length
*	27195	27294:	gap of 100 bp
*	27295	28162:	contig of 868 bp in length
*	28163	28262:	gap of 100 bp
*	28263	29148:	contig of 886 bp in length
*	29149	29248:	gap of 100 bp
*	29249	30154:	contig of 906 bp in length
*	30155	30254:	gap of 100 bp
*	30255	31124:	contig of 870 bp in length
*	31125	31204:	gap of 100 bp
*	31225	32105:	contig of 881 bp in length
*	32106	32205:	gap of 100 bp
*	32206	33093:	contig of 888 bp in length
*	33094	33193:	gap of 100 bp
*	33194	34078:	contig of 885 bp in length
*	34079	34178:	gap of 100 bp

421	CTAATTGGTCTCTTCTCTCGGATCTGGCTTTATCAGATAAAACAAGTGGTCATGCCAC	480
488	AGGATGCTCTATAAGGCCATTGATTCCTGGGATTCATGAGTGATGCTCATATGACTAAGCC	547
481	AGGATGCTCTATAAGCCCATTTGATTCCTGGGATTCATGAGTGATGCTCATATGACTAAGCC	540
548	AGGAGAGACTTATTTAAAGATCTCAGCATCTCTTACCTTGTTAACTAGAGAAACCCGA	607
541	AGGAGAGACTTATTTAAAGATCTCAGCATCTTTTAACTTGTTAACTAGAGAAACCCGA	600
608	AGCATGACTGGATTATTAAGGGGAAATTTGAATCGGTCACCAA	650
601	AGCATGACTGGATTATTAAGGGGAAATTTGAATCGGTCACCAA	643

RESULT 7	AC021520	73806 bp	DNA	linear	HTG 13-JUL-2000
LOCUS	AC021520				
DEFINITION	Homo sapiens chromosome 17 clone RP11-329H16 map 17, LOW-PASS SEQUENCE SAMPLING.				
ACCESSION	AC021520				
VERSION	AC021520.2	GI:9148483			
KEYWORDS	HTG; HTGS PHASE0.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 73806)				
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.				
TITLE	Homo sapiens chromosome 17 clone RP11-329H16				

2 (bases 1 to 73806)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Bock, L., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeFreilano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreirra, P., FitHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lander, T., Lechoczky, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Olivart, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (18-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	On Jul 13, 2000 this sequence version replaced gi:6703580. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <a href="http://www.frc.utah.edu/pw/RepeatMasker.html">http://www.frc.utah.edu/pw/RepeatMasker.html</a>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: L5499  
Center clone name: 329 H 16

\* NOTE: This record contains 76 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of Ns that separate the reads  
\* and the order of which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that the clone



Db	3684	CTTATACCGAAATATTCCCAACACACCCCATATATAAAACTCCCTAACTTCGAAACCCAAAT	3625
Qy	63	ATCTGGAAATCGAGCTCCAGCCAAATGCAATTCCTCTTCTACGGGATCTGGGAATCTCCAAA	122
Db	3624	ATCTAAATAACAATCTCAACCAAAATACATTCTCTTCTACGAAATCTAAAAATCTCCAAA	3565
Qy	123	GCTGCCTCCTCAGATGGGAATTTCCACTCACTTCTCTCAAGCCAGCACTGACCTCCGAG	182
Db	3564	ACTACCTCCTCAAAATAAAAATTTTCCACTCACTTCTCTCAAGCCAACTAAACCTCCCAA	3505
Qy	183	CGGGGAGGGGATCTTTTCTTTGACAGACAGAGTGGGAGGCAGACAGCTGTCACTTTCC	242
Db	3504	CGAAAAAAAACATCTTTTCTTTTAACAAAACAAAATATAAAACAACATATCACTTTCC	3445
Qy	243	AGAAGACTTTCTTTTCTTGATTCATACCCTTTCACCTTCCCTGTGTTTACTGTCTGATATAT	302
Db	3444	AAAAAACTTTTCTTTTCTAAATTCATACCCTTTCACCTTCCCTATATTTTACTATCTAAATAT	3385
Qy	303	GCAAAGGCAAGTCACTTTTCAGAGATGACAACTCTTCTCTGAAGTAGACAGATGCTCC	362
Db	3384	ACAAAAACCAATCACTTTCAAAAATTAACAACCTCTTCTTAAATAAAAAACATATCTCC	3325
Qy	363	AACACTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCT - GGAAGTTTTTCTCTCTGAC	421
Db	3324	AACACTCAAAAACCTATATATAAACHCTCAACCAACAAACTAAAAATTTTCTCTATAAC	3265
Qy	422	CATGGGCTAAATGGTCTCTCTCTGGAATGTGGCTTATCAGATATAAAACAAGTGAGTCA	481
Db	3264	CATAACTAAATTAATCTCTTCTCTAAATATATACTTTATCAATAAAAAACAATAATCA	3205
Qy	482	TGCCACAGAGTGTCTATAAGCCCAATGATCTCGGATTTCTATGATGATCTGATATGAC	541
Db	3204	TACCACAAAATATCTATAAACCCCATTAATCTTAAATTTCTATAAATTAATAATAATAC	3145
Qy	542	TAAAGCCAGGAGAGACTTAATTTAAAGATCTCAGCATCTTTCAAGTTGTTAACCTTAGAGAAA	601
Db	3144	TAAACCAAAAAAAACTTATTTTAAAAATCTCAACATCTTTCAACTTATTAACCTTAAAAAAA	3085
Qy	602	ACCGAAGCATGACTGGATTTATAAAGGAAATTTGATCGGCTCCACCAA	650
Db	3084	ACCCGAAACATACCTAAATATAAAAAAAATTAATACGATCACCA	2036

```

RESULT 9
AX346554
LOCUS          5926 bp      DNA          linear          PAT 01-FEB-2002
DEFINITION     Sequence 1625 from Patent WO200928.
ACCESSION      AX346554
VERSION        AX346554.1  GI:18494440
KEYWORDS       .
SOURCE         synthetic construct
ORGANISM       synthetic construct
               other sequences; artificial sequences.
REFERENCE      1
AUTHORS       Olek A., Piepenbrock, C. and Berlin, K.
TITLE         Diagnosis of diseases associated with the immune system
JOURNAL       Patent: WO 0200928-A 1625 03-JAN-2002;
               EpiGenomics AG (DE)
FEATURES       Location/Qualifiers
               1..5926
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="Chemically treated genomic DNA (Homo sapiens)"

```

[illegible]

Qy	61	GTATCTGGAATGAGGCTCCAGCCAAATGCAATCTCTTCTACGGGATCTGGGAATCTCCA	120
Db	2301	GTATTTGGGAATGAGGTTTGTAAATGATTTTTTTTATACGGGATTTGGGAATTTTGA	2360
Qy	121	AAAGTCGCTCTCAGAGTGGGAATTTCCACTCACTTCTCAGCCAGCACTGACCTCCC	180
Db	2361	AAAGTTGTTTTTAGAGTGGGAATTTTATTTATTTTTTTTACGTTAGTATGATTTTTT	2420
Qy	181	AGGGGGGAGGGGATCTTTCTTGACAGACAGAAAGTGGAGGCGACAGAGCTGTCACTTT	240
Db	2421	AGGGGGGAGGGGTATTTTTTTTTTGATAGAGTAGAAGTGGAGGTAGATGATTTATTTT	2480
Qy	241	CCAGAGACATTTCTTTCTTGATTCATACCTTCACCTCCCTCGTTTACTCTCTGATAT	300
Db	2481	TTAGAGATTTTTTTTTTTTTTGATTAATTTTTTATTTTTTTGTGTTATTTCTTGATAT	2540
Qy	301	ATGCAAAAGGCCAAGTCACTTTTCAGAGATGACAACTCTCTCGAGTAGAGACATGCTT	360
Db	2541	ATGTAAAGGTTAAGTTATTTTTTTTAGAGATGATAATTTTTTTTGAAGTAGAGATATGTT	2600
Qy	361	CCACACTCAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCT--GGAGGTTTTTCTCTGTG	419
Db	2601	TTAATAATTTAGAAGTTTATGTGTAATTTTAGTTAGTAAAGTTGGGAAGTTTTTTTTTGTG	2660
Qy	420	ACCATGGGCTAAATGGTCTCCTTCTCTGGATTTGTGGCTTATCAGATAAAACAAGTGACT	479
Db	2661	ATTATGGGTTAAATGGTTTTTTTTTTTTTTGGATTTGGTTTTTATTTAGATAAAATAAGTGGT	2720
Qy	480	CATGCCACAGGATGCTATAGCCCATGATCTGGGATTTCTATGAGTGAATGCTGATATG	539
Db	2721	TATGTTATAGGATGTTTATAGGTTTATGATTTTTGGGATTTATGAGTGAATGCTGATATG	2780
Qy	540	ACTAAGCCAGGAGAGACTTATTTTAAAGATCTCAGCATCTTTTCACTTGTTTAACCTTAGAG	599
Db	2781	ATTAAATTAGAGAGATTTATTTAAAGATTTTAGTATTTTTTATGTTTGTATTTTAGAGA	2840
Qy	600	AAACCCGAAGCATGACTGGATTATAAGGGGAATTTGATCGGCTCCACCAA	650
Db	2841	AAATTCGAAGTATGATTTGGATTATTAAGGGGAATTTGAATCGGTTTTATTAA	2891

```

RESULT 10
AF493701
LOCUS
DEFINITION
Callithrix jacchus monocyte chemoattractant protein 1 (MCP1) gene,
promoter region.
ACCESSION
AF493701
VERSION
AF493701.1 GI:20530677
KEYWORDS
Callithrix jacchus (white-tufted-ear marmoset)
Callithrix jacchus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
Callithrix.
ORGANISM
1 (bases 1 to 316)
Shanmugasundaram,G.K., Chakraborti,S., Sankaranarayanan,K. and
Banerjee,A.C.
TITLE
Direct Submission
JOURNAL
Submitted (21-MAR-2002) Virology II, National Institute of
Immunology, Aruna Asag Ali Marg, J.N.U. Campus, New Delhi 110 067,
India
FEATURES
source
location/Qualifiers
1..316
/organism="Callithrix jacchus"
/mol_type="genomic DNA"
/db_xref="taxon:9483"
<1..>316
/gene="MCP1"
note="monocyte chemoattractant protein 1"
<1..>316
/gene="MCP1"
gene
promoter
ORIGIN

```

Query Match 22.7%; Score 147.8; DB 9; Length 316;  
Best Local Similarity 85.9%; Pred. No. 2.9e-34;  
Matches 176; Conservative 0; Mismatches 27; Indels 2; Gaps 1;  
QY 328 ATGACAACTCCTTCCTGAAGTAGAGACATGCTTCCAACTCAGAACGCTATGTGAACAC 387  
DB 112 ACGCCAGCACTGACCTCCCGCAGCAGCATGCTTCCAACTCAGAACGCTATGTGAACAC 171  
QY 388 TCAGCCAGCAAAAGCTGG--AAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTC 445  
DB 172 TCAGCCAGCAAAAGCTGGCAAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTC 231  
QY 446 TGGATTGTGGCTTATCAGATAAAACAAGTAGCTATGCCACAGGATGTCTATAAGCCCA 505  
DB 232 TGGATTGTGGCTTATCAGATAAAACAAGTAGCTATGCCACAGGATGTCTATAAGCCCA 291  
QY 506 TTGATTCTGGGATTCATGAGTGAT 530  
DB 292 TTGATTCTGGGATTCATGAGTGAT 316

RESULT 11  
AF493698 307 bp DNA linear PRI 12-MAY-2002  
LOCUS Homo sapiens isolate 2 monocyte chemoattractant protein 1 (MCP1)  
DEFINITION gene, promoter region.  
ACCESSION AF493698  
VERSION AF493698.1 GI:20530674  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 307)  
AUTHORS Sharmugasundaram, G.K., Sankaranarayanan, K. and Banerjee, A.C.  
TITLE Direct Submission  
JOURNAL Submitted (21-MAR-2002) Virology II, National Institute of Immunology, Aruna Asag Ali Marg, J.N.U. Campus, New Delhi 110 067, India

## FEATURES

source Location/Qualifiers  
1..307  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="2"  
/db\_xref="taxon:9606"  
<1..>307  
/gene="MCP1"  
/note="monocyte chemoattractant protein 1"  
promoter <1..>307  
/gene="MCP1"

## ORIGIN

Query Match 21.8%; Score 141.4; DB 9; Length 307;  
Best Local Similarity 90.2%; Pred. No. 2.8e-32;  
Matches 185; Conservative 0; Mismatches 16; Indels 4; Gaps 3;  
QY 328 ATGACAACTCCTTCCTGAAGTAGAGACATGCTTCCAACTCAGAACGCTATGTGAACAC 387  
DB 105 ACGCCAGCACTGACCTCCCGCAGCAGCATGCTTCCAACTCAGAACGCTATGTGAACAC 164  
QY 388 TCAGCCAGCAAAAGCT--GGAAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTCT 446  
DB 165 TCAGCCAGCAAAAGCTGGGAAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTCT 224  
QY 447 GGATTGTGGC--TTATCAGATAAAACAAGTAGCTATGCCACAGGATGTCTATAAGCCCA 505  
DB 225 GGATTGTGGCTTATCAGATAAAACA--GGGTATGCCACAGGATGTCTATAAGCCCA 282  
QY 506 TTGATTCTGGGATTCATGAGTGAT 530  
DB 283 TTGATTCTGGGATTCATGAGTGAT 307

RESULT 12  
AF493700 312 bp DNA linear PRI 12-MAY-2002  
LOCUS Macaca radiata monocyte chemoattractant protein 1 (MCP1) gene,  
DEFINITION promoter region.  
ACCESSION AF493700  
VERSION AF493700.1 GI:20530676  
KEYWORDS  
SOURCE Macaca radiata (bonnet macaque)  
ORGANISM Macaca radiata  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
Cercopitheciinae; Macaca.  
1 (bases 1 to 312)  
AUTHORS Sharmugasundaram, G.K., Chakraborti, S., Sankaranarayanan, K. and Banerjee, A.C.  
TITLE Direct Submission  
JOURNAL Submitted (21-MAR-2002) Virology II, National Institute of Immunology, Aruna Asag Ali Marg, J.N.U. Campus, New Delhi 110 067, India

## FEATURES

source Location/Qualifiers  
1..312  
/organism="Macaca radiata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9548"  
<1..>312  
/gene="MCP1"  
/note="monocyte chemoattractant protein 1"  
promoter <1..>312  
/gene="MCP1"

## ORIGIN

Query Match 21.4%; Score 139; DB 9; Length 312;  
Best Local Similarity 86.3%; Pred. No. 1.6e-31;  
Matches 177; Conservative 0; Mismatches 25; Indels 3; Gaps 2;  
QY 328 ATGACAACTCCTTCCTGAAGTAGAGACATGCTTCCAACTCAGAACGCTATGTGAACAC 387  
DB 109 ACGCCAGCACTGACCTCCCGCAGCAGCATGCTTCCAACTCAGAACGCTATGTGAACAC 168  
QY 388 TCAGCCAGCAAAAGCTGGGAAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTCTG 447  
DB 169 TCAGCCAGCAAAAGCTGGGAAGTTTCTCTGTGACCATGGGCTAATGGTCTCCTTCTCTG 228  
QY 448 GATTGTGGCTT--ATCAGATAAAACAAGTAGCTATGCCACAGGATGTCTATAAGCCCA 505  
DB 229 GATTGTGGCTTGTCTCAGATAAAAGAAGTG--GCCATGCCAGGATGTCTGTAAAGCCCA 287  
QY 506 TTGATTCTGGGATTCATGAGTGAT 530  
DB 288 TTGACTCTGGAATTCATGAGTGAT 312

RESULT 13  
AF493697 310 bp DNA linear PRI 12-MAY-2002  
LOCUS Homo sapiens isolate 1 monocyte chemoattractant protein 1 (MCP1)  
DEFINITION gene, promoter region.  
ACCESSION AF493697  
VERSION AF493697.1 GI:20530673  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 310)  
AUTHORS Sharmugasundaram, G.K., Chakraborti, S., Sankaranarayanan, K. and Banerjee, A.C.  
TITLE Direct Submission  
JOURNAL Submitted (21-MAR-2002) Virology II, National Institute of Immunology, Aruna Asag Ali Marg, J.N.U. Campus, New Delhi 110 067, India

FEATURES	Location/Qualifiers	
	1..310	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/isolate="1"	
	/db_xref="taxon:9606"	
	<1..>310	
	/gene="MCP1"	
	/note="monocyte chemoattractant protein 1"	
	<1..>310	
/gene="MCP1"		
ORIGIN		
Query Match		
Best Local Similarity 21.3%; Score 138.2; DB 9; Length 310;		
Matches 161; Conservative 0; Mismatches 18; Indels 2; Gaps 1;		
QY	350	GAGACATGCTTCCAACTCAGAGCCTATGTGAACACTCAGCCAGCAAGCTGGAAGTT 409
Db	132	GAGACATGCTTCCAACTCAGAGCCTATGTGAACACTCAGCCAGCAAGCTGGAAGTT 191
QY	410	TTTCTCTGACCATGGCTTAATGGTCTCCTTCTCTGGATTGGCTTATCAGATAAAA 469
Db	192	TTTCTCTGACCATGGGTCAACTGGTCTCCCTCTCTGGATTGGCTTATCAGACAAA 251
QY	470	ACAAGTGAGTCATGCCACAGGATGCTATAAGCCCATTTGATTTGGGATTTATGAGTGA 529
Db	252	AGAAGCAGTT--ACACACAGGTCCTCTATACACCCATGATTTCTGGATTCTATGAGTGA 309
QY	530	T 530
Db	310	T 310
RESULT 14		
AF493699 312 bp DNA linear PRI 12-MAY-2002		
LOCUS Papio hamadryas monocyte chemoattractant protein 1 (MCP1) gene, promoter region.		
ACCESSION AF493699.1 GI:20530675		
VERSION AF493699		
KEYWORDS Papio hamadryas (hamadryas baboon)		
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopithecinae; Papio.		
REFERENCE 1 (bases 1 to 312)		
AUTHORS Shanmugasundaram,G.K., Chakraborti,S., Sankaranarayanan,K. and Banerjee,A.C.		
TITLE Direct Submission		
JOURNAL Submitted (21-MAR-2002) Virology II, National Institute of Immunology, Aruna Asag Ali Marg, J.N.U. Campus, New Delhi 110 067, India		
FEATURES	Location/Qualifiers	
	1..312	
source		/organism="Papio hamadryas"
		/mol_type="genomic DNA"
		/db_xref="taxon:9557"
		<1..>312
gene		/gene="MCP1"
		/note="monocyte chemoattractant protein 1"
		<1..>312
promoter		/gene="MCP1"
ORIGIN		
Query Match		
Best Local Similarity 19.4%; Score 126.4; DB 9; Length 312;		
Matches 153; Conservative 0; Mismatches 26; Indels 1; Gaps 1;		
QY	352	GACATGCTTCCAACTCAGAGCCTATGTGAACACTCAGCCAGCAAGCTT 410
Db	133	GACATGCTTCCAACTCAGAGCCTGTCTAAACACTCAGCCAGCAAGCTGGGAAGTTT 192

Db 471 TGGGAAGAACATCCTTTTGTGACAGAGTGAAGTGAGAGGGGAAACAGACAATAACTCTC-- 528  
QY 244 GAAGACTTTCTTTTCTGATTCATACCGCTTCACCTTCCCTGTGTGTTTACTGTCTGATATATG 303  
Db 529 -----CTTTCCTTCGTTTATGATTCACATTTGGCGTGCCTCCTGTGCTAACTGAAGCTTG 583  
QY 304 CAAAGGCCAAGTCACCTTTCCAGAGATGA 331  
Db 584 CAGCGGCCAATTCACCTTTGCAATGGTGA 611

Search completed: August 4, 2005, 11:17:04  
Job time : 2616.63 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 430.789 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-7  
Perfect score: 650  
Sequence: 1 ggctgtgctgcgagatgtcc.....aattgaatgcggtccaccaa 650

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001s:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	650	100.0	650	ADH13944	Adh13944 Human mon
2	616.6	94.9	11793	12 AD003803	Ado03803 Human Ccl
3	608.6	93.6	3221	12 ADN12120	Adn12120 MCP1 gene
4	608.6	93.6	3221	12 AD003802	Ado03802 Human Ccl
5	513.2	79.0	599	8 ACF64397	Acf64397 Human MCP
6	401.8	61.8	5926	6 ABL33653	Ab133653 Human imm
7	365.4	56.2	5926	6 ABL33652	Ab133652 Human imm
8	43.2	6.6	2000	8 ADA71938	Ada71938 Rice gene
9	38.2	5.9	6557	6 ABL33000	Ab133000 Human imm
10	36.6	5.6	1497	10 ADB31378	Adb31378 Testoster
11	35.8	5.5	34796	13 ACN37240	Acn37240 Human per
12	35.8	5.5	211257	11 ACN44200	Acn44200 Mouse gen
13	35.4	5.4	110000	12 ADQ97331_0	Adq97331 Mouse can
14	35.2	5.4	294	4 AAI91704	Aai91704 Human pol
15	34.8	5.4	99886	11 ACN44066	Acn44066 Human gen
16	34.8	5.4	192992	13 ABL32866	Abd32866 Mouse can
17	34.6	5.3	19576	6 ABL70575	Ab170575 Chemical
18	34.6	5.3	19576	6 AAS61258	Aas61258 Human gen
19	34.4	5.3	1947	6 ABE212660	Abz12660 Arabidops
20	34.4	5.3	6171	6 ABL33011	Ab133011 Human imm

21	34	5.2	402	5	AAF65257	Aaf65257 Novel hum
22	34	5.2	682	5	AAF65185	Aaf65185 Novel hum
23	34	5.2	2243	11	ADM01616	Adm01616 Human cDN
24	33.8	5.2	502	6	ABN73535	Abn73535 Bovine em
25	33.8	5.2	2000	8	ADA71938	Ada71938 Rice gene
26	33.8	5.2	6759	6	ABL32738	Ab132738 Human imm
27	33.8	5.2	15923	6	ABK31220	Abk31220 Signal tr
28	33.8	5.2	15923	6	ABL70179	Ab170179 Chemical
29	33.8	5.2	15923	6	AAS61132	Aas61132 Human gen
30	33.8	5.2	24259	4	AAS46692	Aas46692 Tumour su
31	33.6	5.2	885	6	ABZ13218	Abz13218 Arabidops
32	33.6	5.2	16509	6	ABL33321	Ab133321 Human imm
33	33.6	5.2	24792	13	ABD33027	Abd33027 Human can
34	33.6	5.2	34548	6	AB170604	Ab170604 Chemical
35	33.6	5.2	50000	3	AAA96366	Aaa96366 Polymorph
36	33.4	5.1	1439	4	AAI84699	Aai84699 Human pol
37	33.4	5.1	1759	12	ADL12544	Adl12544 Human ste
38	33.4	5.1	10757	4	ABL21478	Ab121478 Drosophil
39	33.4	5.1	14279	10	ADB52913	Adb52913 Primary r
40	33.4	5.1	110000	2	AAV21209_02	Continuation (3 of
41	33.2	5.1	1144	3	AAC56023	Aac56023 Eucalyptu
42	33.2	5.1	110000	6	ABA90193_0	Abas90193 Human oes
43	33.2	5.1	110000	6	ABQ87681_0	Abq87681 Human oes
44	33.2	5.1	110000	8	ABX33717_0	Abx33717 Gene enco
45	33.2	5.1	110000	10	ADB81391_0	Adb81391 Partial g

ALIGNMENTS

RESULT 1

ADH13944  
ID ADH13944 standard; DNA; 650 BP.  
XX  
AC ADH13944;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:7.  
XX  
KW ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
KW cervical carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN EP1170372-A1.  
XX  
PD 09-JAN-2002.  
XX  
PF 06-JUL-2000; 2000EP-00114560.  
XX  
PR 06-JUL-2000; 2000EP-00114560.  
XX  
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
PI Roessl F, Soto U, Coy J, Finzer P, Delius H, Poustka A;  
PI Zur Hausen H, Patzelt A;  
XX  
WPI; 2002-165895/22.  
PT Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
PT useful in gene therapy, for treating atherosclerosis and cancer.  
XX  
XX Disclosure; SEQ ID NO 7; 30pp; English.  
XX  
CC The invention relates to a novel nucleic acid molecule (I) comprising a  
CC sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
CC having the biological activity of MCP-1. A protein encoded by a nucleic  
CC acid of the invention has cytostatic, and antiarteriosclerotic activity.  
CC A nucleic acid of the invention may have a use in gene therapy. A  
CC compound of the invention is useful in the preparation of a medicament  
CC for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

CC	pharmaceutical composition of the invention is useful for the treatment	PN	WO2004041160-A2.
CC	of diseases associated with dysregulation of MCP-1 expression, e.g.	XX	
CC	atherosclerosis or cancer. The present sequence is used in the	PD	21-MAY-2004.
CC	exemplification of the invention.	XX	
XX		PF	16-OCT-2003; 2003WO-US032933.
SQ	Sequence 650 BP; 167 A; 165 C; 147 G; 171 T; 0 U; 0 Other;	PR	30-OCT-2002; 2002US-0422096P.
		XX	(KENT ) UNIV KENTUCKY RES FOUND.
		XX	Ambati J;
		XX	WPI; 2004-400512/37.
		XX	Testing candidate drug for treating age-related macular degeneration, by
		XX	administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and
		PT	analyzing development or regression of drusen and/or lipofuscin
		PT	accumulation in eye.
		XX	Disclosure; SEQ ID NO 4; 64pp; English.
		PS	This invention relates to a novel methods and animal models for testing
		XX	candidate drugs that can be used for the treatment or prevention of age-
		CC	related macular degeneration (AMD). Specifically, it refers to
		CC	administering a candidate drug to gene knockout mice, in particular Ccl2-
		CC	deficient, Ccr2-deficient and/ or a Ccl2-deficient/Ccr2-deficient dual
		CC	knockout mouse. The present invention describes analysing the knockout
		CC	mouse eye for development or regression of drusen and/ or lipofuscin
		CC	accumulation, as well as for the effect of a candidate drug on Bruch's
		CC	membrane, retinal degeneration and/ or choroidal neovascularisation.
		CC	Accordingly, such compositions exhibit ophthalmological activities and
		CC	can be used for gene therapy purposes. This polynucleotide sequence is
		CC	the human Ccl2 gene and enhancer region DNA of the invention.
		XX	Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;
		SQ	
			Query Match 94.9%; Score 616.6; DB 12; Length 11793;
			Best Local Similarity 97.7%; Pred. No. 1.1e-202;
			Matches 636; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
QY	1 GCGTTGTGCGAGATGTTCCCGACAGCCGATGTGAGAGCTCCCTGGCTCCGGGCCCA 60	QY	1 GCGTTGTGCGAGATGTTCCCGACAGCCGATGTGAGAGCTCCCTGGCTCCGGGCCCA 60
Db	1 GCGTTGTGCGAGATGTTCCCGACAGCCGATGTGAGAGCTCCCTGGCTCCGGGCCCA 60	Db	4797 GCGTTGTGCGAGATGTTCCCGACAGCCGATGTGAGAGCTCCCTGGCTCCGGGCCCA 4856
QY	61 GTATCTGGATGAGCTCCAGCCAAATGCTTCTCTACGGGATCTGGAACTTCCA 120	QY	61 GTATCTGGATGAGCTCCAGCCAAATGCTTCTCTACGGGATCTGGAACTTCCA 120
Db	61 GTATCTGGATGAGCTCCAGCCAAATGCTTCTCTACGGGATCTGGAACTTCCA 120	Db	4857 GTATCTGGATGAGCTCCAGCCAAATGCTTCTCTACGGGATCTGGAACTTCCA 4916
QY	121 AAGCTGCTCTCTACGGGATCTGGAACTTCCA 180	QY	121 AAGCTGCTCTCTACGGGATCTGGAACTTCCA 180
Db	121 AAGCTGCTCTCTACGGGATCTGGAACTTCCA 180	Db	4917 AAGCTGCTCTCTACGGGATCTGGAACTTCCA 4976
QY	181 AGCGGGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 240	QY	181 AGCGGGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 240
Db	181 AGCGGGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 240	Db	4977 AGCGGGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 5036
QY	241 CCAGAGACTTTCTTTCTGATTCATACCTTCACTTCTCTACGGGATCTGGAACTTCCA 300	QY	241 CCAGAGACTTTCTTTCTGATTCATACCTTCACTTCTCTACGGGATCTGGAACTTCCA 300
Db	241 CCAGAGACTTTCTTTCTGATTCATACCTTCACTTCTCTACGGGATCTGGAACTTCCA 300	Db	5037 CCAGAGACTTTCTTTCTGATTCATACCTTCACTTCTCTACGGGATCTGGAACTTCCA 5096
QY	301 ATGCAAGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 360	QY	301 ATGCAAGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 360
Db	301 ATGCAAGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 360	Db	5097 ATGCAAGGAGGAGCTCTTTTCTGAGATGACAACTCTCTCTACGGGATCTGGAACTTCCA 5156
QY	361 CCAACTCAGAGGCTATGTGAACACTCAGCAGCAAGCT-GGAAGTTTTCTCTGTG 419	QY	361 CCAACTCAGAGGCTATGTGAACACTCAGCAGCAAGCT-GGAAGTTTTCTCTGTG 419
Db	361 CCAACTCAGAGGCTATGTGAACACTCAGCAGCAAGCT-GGAAGTTTTCTCTGTG 419	Db	5157 CCAACTCAGAGGCTATGTGAACACTCAGCAGCAAGCT-GGAAGTTTTCTCTGTG 5216
QY	420 ACCATGGCTAATGGTCTCTCTCTGATTTGGCTTATCAGATATAAACAAGTGAGT 479	QY	420 ACCATGGCTAATGGTCTCTCTCTGATTTGGCTTATCAGATATAAACAAGTGAGT 479
Db	420 ACCATGGCTAATGGTCTCTCTCTGATTTGGCTTATCAGATATAAACAAGTGAGT 479	Db	5217 ACCATGGCTAATGGTCTCTCTCTGATTTGGCTTATCAGATATAAACAAGTGAGT 5276
QY	480 CATGCCACAGGATGTCTATAAGCCCAATGATTCTGGGATTTATGAGTGTGATGATG 539	QY	480 CATGCCACAGGATGTCTATAAGCCCAATGATTCTGGGATTTATGAGTGTGATGATG 539
			RESULT 2
			AD003803
			ID AD003803 standard; DNA; 11793 BP.
			AC AD003803;
			XX AD003803;
			DT 12-AUG-2004 (first entry)
			Human Ccl2 gene and enhancer region DNA SeqID 4.
			DE human; ds; animal model; age-related macular degeneration; AMD;
			XX gene knockout; Ccl2-deficient; Ccr2-deficient; drusen;
			KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;
			KW choroidal neovascularisation; ophthalmological; gene therapy.
			XX Homo sapiens.
			OS
			XX

```
Db 5277 CATGCCACAGGATGCTATAAGCCCATTTGATTCTGGGATTTCTATGATGCTGATATG 5336
QY 540 ACTAAGCCAGGAGACTTATTTAAGATCTCAGCATCTTTTCACTGTTTAACTTAGAGA 599
Db 5337 ACTAAGCCAGGAGAGACTTATTTAAGATCTCAGCATCTTTTCACTGTTTAACTTAGAGA 5396
QY 600 AAACCCGAAGCATGACTGGATTATAAGGGGAAATTTGAATGCGGTCCACCAA 650
Db 5397 AAACCCGAAGCATGACTGGATTATAAGGGGAAATTTGAATGCGGTCCACCAA 5447

RESULT 3
ADN12120
ID ADN12120 standard; DNA; 3221 BP.
XX AC
XX AC
XX AC
XX DT 17-JUN-2004 (first entry)
XX DT MCP1 gene promoter region.
XX DE
XX DE
XX KW major histocompatibility class I; MHC-I; MHC-II; Cytostatic;
XX KW EBV-associated cancer; Hodgkin's lymphoma; nasopharyngeal carcinoma;
XX KW gastric carcinoma; Burkitt's lymphoma; T-cell lymphoma; B-cell lymphoma;
XX KW parotid carcinoma; breast carcinoma; leiomyosarcoma; ds.
XX OS
XX OS Homo sapiens.
XX PN WO2004027036-A2.
XX PD
XX PD 01-APR-2004.
XX PF 19-SEP-2003; 2003WO-US029684.
XX PF 19-SEP-2002; 2002US-0411990P.
XX PR
XX PR (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX PA
XX PA Ambinder RF, Yang Y, Borrello IM, Levitsky HI;
XX WPI; 2004-295406/27.
XX DR
XX PT New human cell line modified to comprise and express genes encoding
XX PT immunomodulators and an antigen of Epstein-Barr virus (EBV), useful for
XX PT inducing or stimulating an immune response in a human to EBV-associated
XX PT cancer.
XX PS Disclosure; SEQ ID NO 13; 218pp; English.
XX CC The present invention relates to a human cell line, which lacks major
XX CC histocompatibility class I (MHC-I) and MHC-II antigens and which has been
XX CC modified to comprise and express a gene encoding an immunomodulator and a
XX CC gene encoding an antigen of Epstein-Barr virus (EBV). The human cell
XX CC line, compositions and methods are useful for inducing or stimulating an
XX CC immune response in a human to an EBV-associated cancer, where the human
XX CC has or is at risk for Hodgkin's lymphoma, nasopharyngeal carcinoma,
XX CC gastric carcinoma, Burkitt's lymphoma, T-cell lymphoma, B-cell lymphoma,
XX CC parotid carcinoma, breast carcinoma, and leiomyosarcoma. The present
XX CC sequence represents a nucleotide sequence associated with the cell line
XX CC of the invention.
XX SQ Sequence 3221 BP; 859 A; 793 C; 727 G; 842 T; 0 U; 0 Other;

Query Match 93.8%; Score 608.6; DB 12; Length 3221;
Best Local Similarity 97.7%; Pred. No. 3.2e-200;
Matches 628; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 9 CCGAGATGTTCCAGACACAGCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG 68
Db 1 CCGAGATGTTCCAGACACAGCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG 60
QY 69 AATGCAGGCTCCAGCCAAATGCAATTTCTTCTACGGGATCTGGGAATCTCCAAAGCTGCC 128
```

```
Db 61 AATGCAGGCTCCAGCCAAATGCAATTTCTTCTACGGGATCTGGGAATCTCCAAAGCTGCC 120
QY 129 TCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGG 188
Db 121 TCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGG 180
QY 189 AGGGCATCTTTTCTTGACAGAGCAGAAGTGGGAGGCGAGACAGCTGTCACCTTTCCAGAAGA 248
Db 181 AGGGCATCTTTTCTTGACAGAGCAGAAGTGGGAGGCGAGACAGCTGTCACCTTTCCAGAAGA 240
QY 249 CTTTCTTTTCTGATTCATACCTTTCACTTCCCTCTGTGTTTACTGTCTGATATATGCAAG 308
Db 241 CTTTCTTTTCTGATTCATACCTTTCACTTCCCTCTGTGTTTACTGTCTGATATATGCAAG 300
QY 309 GCCAAGTCATTTTCCAGAGATGACAACTCTTCTGGAAGTAGAGACATCTTCCACACT 368
Db 301 GCCAAGTCATTTTCCAGAGATGACAACTCTTCTGGAAGTAGAGACATCTTCCACACT 360
QY 369 CAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCT-GGAAGTTTTTCTCTGTGACCATGGG 427
Db 361 CAGAAGCCTATGTGAACACTCAGCCAGCAAAAGCTGGAAGTTTTTCTCTGTGACCATGGG 420
QY 428 CTAATTTGGTCTCCTTCTCTGGAATGTGGCTTATCAGATAAAAAACAAGTGAGTCATGCCAC 487
Db 421 CTAATTTGGTCTCCTTCTCTGGAATGTGGCTTATCAGATAAAAAACAAGTGAGTCATGCCAC 480
QY 488 AGGATGTCTATAAGCCCAATTGATTTCTGGGATCTATGATGATGCTGATATGACCTAAGCC 547
Db 481 AGGATGTCTATAAGCCCAATTGATTTCTGGGATCTATGATGATGCTGATATGACCTAAGCC 540
QY 548 AGGAGAGACTTATTTAAAGATCTCAGCATCTTTTCAAGCTTTTAACTAGAGAAAACCCGA 607
Db 541 AGGAGAGACTTATTTAAAGATCTCAGCATCTTTTCAAGCTTTTAACTAGAGAAAACCCGA 600
QY 608 AGCATGACTGGATTATAAGGGGAAATTTGAATCGGTCCACCAA 650
Db 601 AGCATGACTGGATTATAAGGGGAAATTTGAATCGGTCCACCAA 643

RESULT 4
AD003802
ID AD003802 standard; DNA; 3221 BP.
XX AC
XX AC AD003802;
XX DT 12-AUG-2004 (first entry)
XX DE
XX DE Human Ccl2 promoter region used for gene knockout animal models SeqID 3.
XX KW human; ds; animal model; age-related macular degeneration; AMD;
XX KW gene knockout; Ccl2-deficient; Ccr2-deficient; drusen;
XX KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;
XX KW choroidal neovascularisation; ophthalmological; gene therapy; promoter.
XX OS
XX OS Homo sapiens.
XX PN WO2004041160-A2.
XX XX
XX PD 21-MAY-2004.
XX PF 16-OCT-2003; 2003WO-US032933.
XX PR 30-OCT-2002; 2002US-0422096P.
XX PA (KENT ) UNIV KENTUCKY RES FOUND.
XX PI Ambati J;
XX DR WPI; 2004-400512/37.
XX PT Testing candidate drug for treating age-related macular degeneration, by
XX PT administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and
```

PT analyzing development or regression of drusen and/or lipofuscin  
PT accumulation in eye.

XX Disclosure; SEQ ID NO 3; 64pp; English.

XX This invention relates to a novel methods and animal models for testing  
CC candidate drugs that can be used for the treatment or prevention of age-  
CC related macular degeneration (AMD). Specifically, it refers to  
CC administering a candidate drug to gene knockout mice, in particular Ccl2-  
CC deficient, Ccr2-deficient and/ or a Ccl2-deficient/Ccr2-deficient dual  
CC knockout mouse. The present invention describes analysing the knockout  
CC mouse eye for development or regression of drusen and/ or lipofuscin  
CC accumulation, as well as for the effect of a candidate drug on Bruch's  
CC membrane, retinal degeneration and/ or choroidal neovascularisation.  
CC Accordingly, such compositions exhibit ophthalmological activities and  
CC can be used for gene therapy purposes. This polynucleotide sequence is  
CC the human Ccl2 promoter region DNA of the invention.

XX SQ Sequence 3221 BP; 859 A; 793 C; 727 G; 842 T; 0 U; 0 Other;

Query Match 93.6%; Score 608.6; DB 12; Length 3221;  
Best Local Similarity 97.7%; Pred. No. 3.2e-200;  
Matches 628; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 9 CCGAGATGTTCCAGACAGCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG 68  
DB 1 CCGAGATGTTCCAGACAGCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGTATCTGG 60  
QY 69 AATGCAGGCTCCAGCCAAATGCAATCTCTTACGGGATCTGGGAATTCCTCAAGAGTCC 128  
DB 61 AATGCAGGCTCCAGCCAAATGCAATCTCTTACGGGATCTGGGAATTCCTCAAGAGTCC 120  
QY 129 TCCTCAGAGTGGGAATTCACATCTCTCTCAAGAGTCTGGGAATTCCTCAAGAGTCC 188  
DB 121 TCCTCAGAGTGGGAATTCACATCTCTCTCAAGAGTCTGGGAATTCCTCAAGAGTCC 180  
QY 189 AGGCATCTTTCTTGACAGACAGAGTGGGAGGAGCAGAGTGTCACTTTCCAGAGA 248  
DB 181 AGGCATCTTTCTTGACAGACAGAGTGGGAGGAGCAGAGTGTCACTTTCCAGAGA 240  
QY 249 CTTTCTTTCTGATTATACCTTTCACTTCCCTGTGTCTTCTGTATATATGCAAG 308  
DB 241 CTTTCTTTCTGATTATACCTTTCACTTCCCTGTGTCTTCTGTATATATGCAAG 300  
QY 309 GCCAAGTCACTTCCAGAGATGACAACTCTTCTCTGAGTACAGATGCTTCCAACT 368  
DB 301 GCCAAGTCACTTCCAGAGATGACAACTCTTCTCTGAGTACAGATGCTTCCAACT 360  
QY 369 CAGAAGCTATGTGAACACTCAGCAGCAAGAGCT-GGAAGTTTTTCTCTGTGACCATGGG 427  
DB 361 CAGAAGCTATGTGAACACTCAGCAGCAAGAGCTGGGAAGTTTTTCTCTGTGACCATGGG 420  
QY 428 CTAATTTGGTCTCTCTCTGTGATTTGGCTTATCAGATAAAACAAGTGAATGATGCCAC 487  
DB 421 CTAATTTGGTCTCTCTCTGTGATTTGGCTTATCAGATAAAACAAGTGAATGATGCCAC 480  
QY 488 AGGATGCTATAAGCCCATTAATTTCTGGATTTCTATGAGTATGCTCATATGACTAAGCC 547  
DB 481 AGGATGCTATAAGCCCATTAATTTCTGGATTTCTATGAGTATGCTCATATGACTAAGCC 540  
QY 548 AGGAGAGACTTATTTAAAGATCTCAGACATCTTTCAAGCTTTGTAACCTAGAGAAAACCCGA 607  
DB 541 AGGAGAGACTTATTTAAAGATCTCAGACATCTTTCAAGCTTTGTAACCTAGAGAAAACCCGA 600  
QY 608 AGCATGACTGGATTATAAGGGAATTTGAATTCGGTCCACCA 650  
DB 601 AGCATGACTGGATTATAAGGGAATTTGAATTCGGTCCACCA 643

RESULT 5  
ACF64397  
ID ACF64397 standard; DNA; 599 BP.  
XX

AC ACF64397;  
XX 13-OCT-2003 (first entry)  
DT  
XX Human MCP1 nucleotide sequence >MCP1\_pro.  
DE  
XX Human; detection; computer-readable storage medium; polymorphic site;  
KW signal carrying data; data processing system; multiple sclerosis; gene;  
KW ds.  
XX Homo sapiens.  
XX WO2003014319-A2.  
XX 20-FEB-2003.  
XX 07-AUG-2002; 2002WO-US025268.  
XX 07-AUG-2001; 2001US-0310741P.  
XX 24-SEP-2001; 2001US-0324790P.  
XX (DNAS-) DNA SCI INC.  
XX Jones HB, Xu H, White R, Rienhoff HY, Jin W, Natsoulis G;  
XX WPI; 2003-268196/26.  
XX New polynucleotide, useful for detecting loci associated with multiple  
XX sclerosis.  
XX Disclosure; Page 74; 93pp; English.  
XX The present invention describes an isolated polynucleotide (PN)  
CC comprising: (a) a sequence comprising at least 15 contiguous nucleotides  
CC of a sequence comprising variant sequences (A) from Table 4 given in the  
CC specification; or (b) a sequence that is complementary to (A). Also  
CC described: (1) an array of (PN)s comprising two or more of the isolated  
CC (PN)s; (2) detecting a (PN) in an individual; (3) a computer-readable  
CC storage medium, where each record has a field identifying a base  
CC occupying a (PN) site and a location of the polymorphic site; and (4) a  
CC signal carrying data for access by an application program having executed  
CC on a data processing system. The (PN) can be used for detecting loci  
CC associated with multiple sclerosis. ACF64025 to ACF64424 represent  
CC sequences used in the exemplification of the present invention  
XX  
XX Sequence 599 BP; 145 A; 160 C; 139 G; 154 T; 0 U; 1 Other;  
SQ  
Query Match 79.0%; Score 513.2; DB 8; Length 599;  
Best Local Similarity 97.0%; Pred. No. 2e-167;  
Matches 543; Conservative 1; Mismatches 14; Indels 2; Gaps 2;  
QY 1 GGCTTTGTGCGGAGATGTTCCAGCACAGCCCATGTGAGAGTCCCTGGCTCCGGGCCCA 60  
DB 41 GGCTTTGTGCGGAGATGTTCCAGCACAGCCCATGTGAGAGTCCCTGGCTCCGGGCCCA 100  
QY 61 GTATCTGGAATCAGGCTCCAGCAATGCAATCTCTTCTAGGGATCTGGGAATCTCA 120  
DB 101 GTATCTGGAATCAGGCTCCAGCAATGCAATCTCTTCTACGGATCTGGGAATCTCA 160  
QY 121 AAGTCTCCTCTCAGAGTGGGAATTTCCACTCACTTCTCTACGCCAGCACTGACTCC 180  
DB 161 AAGTCTCCTCTCAGAGTGGGAATTTCCACTCACTTCTCTACGCCAGCACTGACTCC 220  
QY 181 AGCGGGGAGGGGATCTTTTCTTGACAGACAGAGTGGGAGGAGAGCTGTCACTTT 240  
DB 221 AGCGGGGAGGGGATCTTTTCTTGACAGACAGAGTGGGAGGAGAGCTGTCACTTT 280  
QY 241 CCAGAGAGCTTTCTTTTCTGATTACATCCCTTCCCTTCTGTGTCTGTCTGTATAT 300  
DB 281 CCAGAGAGCTTTCTTTTCTGATTACATCCCTTCCCTTCTGTGTGTGTGTGTGTATAT 340  
QY 301 ATGCAAGGCCAAGTCACTTTCCAGAGATGACAACTCTCTCTGAAGTAGAGATGCTTT 360

Db 341 A-GCAAGGCCAAGTCACCTTCCAGAGATGACAACTCTCTTCTGAAAGTAGACATGCTT 399  
QY 361 CCAACTCTAGAAGCCTATGTGAACACTCAGCCAGCAAGCT-GGAAGTTTTTCTCTGTG 419  
Db 400 CCAACTCTAGAAGCCTATGTGAACACTCAGCCAGCAAGCTGGAAGTTTTTCTCTGTG 459  
QY 420 ACCATGGCTTAATGGTCTCTCTCTCTGGATTGGCTTATCAGATAAAACAAGTGAGT 479  
Db 460 ACCATGGCTTAATGGTCTCTCTCTCTGGATTGGCTTATCAGATAAAACAAGTGGT 519  
QY 480 CATGCCACAGGATGCTATAGGCCCATTTGATCTGGATCTATGATGATGCTGATATG 539  
Db 520 CATGCCACAGGATGCTATAGGCCCATTTGATCTGGATCTATGATGATGCTGATATG 579  
QY 540 ACTAAGCCAGGAGACTTA 559  
Db 580 ACTAAGCCAGGAGACTTA 599

RESULT 6  
ID ABL33653/c  
XX ABL33653 standard; DNA; 5926 BP.  
AC ABL33653;  
XX 26-MAR-2002 (first entry)  
XX Human immune system associated gene SEQ ID NO: 1626.  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
ds.  
XX Homo sapiens.  
XX WO200200928-A2.  
XX 03-JAN-2002.  
XX 02-JUL-2001; 2001WO-EP007537.  
XX PF  
XX 30-JUN-2000; 2000BE-01032529.  
XX PR  
XX 01-SEP-2000; 2000DE-01043826.  
XX  
XX (BPIG-) EPIGENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-130909/17.  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
XX Claim 1; SEQ ID NO 1626; 32pp + Sequence Listing; German.  
PS  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
SQ Sequence 5926 BP; 1554 A; 55 C; 1492 G; 2825 T; 0 U; 0 Other;

Query Match 61.8%; Score 401.8; DB 6; Length 5926;

Best Local Similarity 77.2%; Pred. No. 4.9e-128;  
Matches 501; Conservative 0; Mismatches 147; Indels 1; Gaps 1;  
QY 3 CTGTGCGCAGATGTTCCAGCAGACAGCCCATGTGAGAGCTCCCTGGCTCCGGGCCAGT 62  
Db 3684 CTTATACGAAATATTTCCCAACACACACCCCATATAAAAACTCCCTAACTCCGAACCAAT 3625  
QY 63 ATCTGGAATGCAAGGCTCCAGCCAAATGCAATCTCTTCTACGGGATCTGGGAATCTCCAAA 122  
Db 3624 ATCTAAATACAACTCCCAACCAATACATTTCTTCTACGAAATCTAAAACTTCCAAA 3565  
QY 123 GTGCTCTCTCAGAGTGGGAATTTCCACTCTCTCTCACGCCAGACTGACCTCCCAAG 182  
Db 3564 ACTACCTCTCAAAATATAAAATTTCCACTCTCTCTCACGCCACACTAACCTCCCAA 3505  
QY 183 CGGGGGAGGGCATCTTTCTGACAGAGAGAGTGGGAGGAGAGCTGTCACATTTCC 242  
Db 3504 CGAAAAAAACATCTTTCTTAACAAAAAATAAAAAAACAACATATACATTTCC 3445  
QY 243 AGAAGACTTTCTTTCTGATTCTATACCTTTCACTCTTCTGTTTACTGTCTGATATAT 302  
Db 3444 AAAAATCTTTCTTAATTCATACCTTTCACTCTTCTGTTTACTGTCTGATATAT 3385  
QY 303 GCAAGGCCAAGTCTATTTCCAGAGATGACAACTCTCTCTCTGAAAGTAGAGATGCTTCC 362  
Db 3384 ACAAAAAACCAATCACTTTCCAAAAATAACAACTCTCTCTAAATAAAAAACATATCTCC 3325  
QY 363 AACACTCAGAAGCCTATGACACTCAGCCAGCAAGCT-GGAAGTTTTTCTCTGTGAC 421  
Db 3324 AACACTCAAAAAACCTATATAAACACTCAACCAAAAACTAAAAAATTTTTCTCTATAAC 3265  
QY 422 CATGGCTAAATGGTCTCTCTCTCTGGATTGGGCTTATCAGATAAAAAACAAGTAGTCA 481  
Db 3264 CATAACTAAATTAATCT 3205  
QY 482 TGCCACAGAGATGCTATAGCCCATGATCTCTGGATCTCTAGAGTGTCTGATATGAC 541  
Db 3204 TACCACAAAAATATCTATAAACCCATTAATTTCTAAAAATCTATAATAATACTATAATAAC 3145  
QY 542 TAAGCCAGGAGAGACTTATTTAAAGTCTCAGACTCTTTCAGCTTGTAACTAGAGAA 601  
Db 3144 TAAACCAAAAAAACAATTTAAAAAATCTCAACATCTTTCAACTTATTAACTAAAAAA 3085  
QY 602 ACCCGAAGCATGCTGGATTATAAGGGAATTTGAATGGGTCCACCAA 650  
Db 3084 ACCCGAAGCATGCTGGATTATAAGGGAATTTGAATGGGTCCACCAA 3036

RESULT 7  
ID ABL33652  
XX ABL33652 standard; DNA; 5926 BP.  
AC ABL33652;  
XX  
XX 26-MAR-2002 (first entry)  
XX Human immune system associated gene SEQ ID NO: 1625.  
XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytosine methylation; antiasthmatic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
ds.  
XX Homo sapiens.  
XX WO200200928-A2.  
XX 03-JAN-2002.  
XX

2841 AAATTCCGAAGTATGATTGGATTATTAAGGGAATTAAGTTCGGTATTATTA 2891

Db

RESULT 8  
ADA71938  
ID ADA71938 standard; DNA; 2000 BP.

XX AC ADA71938;  
XX DT 20-NOV-2003 (first entry)  
XX DE Rice gene, SEQ ID 5263.  
XX KW Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; ds.  
XX OS Oryza sativa.  
XX FN WO2003000898-A1.  
XX PD 03-JAN-2003.  
XX PF 22-JUN-2001; 2001WO-IB0011105.  
XX PR 22-JUN-2001; 2001WO-IB0011105.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX CH Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX ID Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
PS Claim 27; SEQ ID NO 5263; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
Query Match 6.6%; Score 43.2; DB 8; Length 2000;  
Best Local Similarity 11.2%; Pred No. 0.002;  
Matches 35; Conservative 154; Mismatches 122; Indels 1; Gaps 1

QY 309 GCCAAGTCACCTTCCAGAGATGACAACTCCCTTCGAGTAGAGACATGCTCCACACT 368  
DB 533 SYSAAKARCVRGKGYWAGMMWKRYKRYMYKMMWYKRYKSCWYCKMSYATSCWK 592  
QY 369 CAGAAGCCTATGACACTCAGCCAGCAAAAGCTGGAAGTCTTCTCTGTGACCATGG-G 427  
DB 593 SARAKAGKCKRSKMSAWSKMSRSKCKCASKRSKRYAMMGMTSGSRMSRWKSYT 652  
QY 428 CTAATTTGGTCTCCTCTCTGGATTTGGCTTATCAGATAAAACAAGTGAATGCCAC 487  
DB 653 CYWRKWSMKSTCTWYMYNSKYTYAKYGVYRYRYAWCMWRWYFYRYRYTWTYAWY 712  
QY 488 AGAATGCTATAAGCCCATGATTCTCGGGATCTCTAGTAGTGATGCTGATGACTAAGCC 547  
DB 713 TSSTRMAMTKMSYGRYTSWYKCYKSKWYRSWYYSWWAKTWKWRFYATRMWMM 772

PF	02-JUL-2001; 2001WO-EP007537.
XX	
PR	30-JUN-2000; 2000DE-01032529.
XX	01-SEP-2000; 2000DE-01043826.
XX	(EPIG-) EPIGENOMICS AG.
PA	Olek A, Piepenbrock C, Berlin K;
PI	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful for
PT	diagnosis and treatment of diseases associated with abnormal cytosine
PT	methylation.
XX	
XX	Claim 1; SEQ ID NO 1625; 32pp + Sequence Listing; German.
XX	
XX	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention
XX	
SQ	Sequence 5926 BP; 1562 A; 55 C; 1318 G; 2991 T; 0 U; 0 Other;
Query Match            56.2%; Score 365.4; DB 6; Length 5926;	
Best Local Similarity   73.6%; Pred. No. 2.3e-115;	
Matches 479; Conservative 0; Mismatches 171; Indels 1; Gaps 1;	
QY	1 GGCTTGTGCCGAGATGTTCCACGACAGCCCCCATGTGAGAGTCCTCGCTCCGGCCCA 60
Db	2241 GGTTTGTCGAGATGTTTTTAGTATAGTTTTATGTGAGAGTTTTTTTGGTTTTCGGTTTTA 2300
QY	61 GTATCTGGAAATCAGGCTCCAGCCAATGCATTCTCTCTACGGGATCTGGGAATCTCCA 120
Db	2301 GTATTTGGAAATGTAGGTTTTTAGTTAAATGTAATTTTTTTTACGGGAATTTGGGAATTTTA 2360
QY	121 AAGCTGCTCTCTAGAGTGGGAATTTCCACTCACTCTCTCACGCCAGCACTGAOCTCCC 180
Db	2361 AAGTTGTTTTTTAGATGGGAATTTTATTTATTTTTTTTACGTTAGTATTGATTTTTT 2420
QY	181 AGCGGGGAGGCAATCTTTTCTTGACAGACAGAAGTGGGAGCAGACAGCTGTCACTTT 240
Db	2421 AGCGGGGAGGGTATTTTTTTTGTATAGAGTAGAAGTGGGAGTAGATAGTTATTAATTT 2480
QY	241 CCAGAAGACTTTCTTTTCTGATTCATAACCCTTCACTTCCCTGTGTGTTTACTGTCTCATAT 300
Db	2481 TTAGAAGATTTTTTTTTTTTGATTTATTTATTTTTTTTTTTTGTTTTGTTTTCATAT 2540
QY	301 ATGCAAAGCCCAAGTCACTTTTCCAGAGATGACAATCTCTTCTGTAAGTAGACATGCTT 360
Db	2541 ATGTAAGGTTAAAGTTATTTTTTTTAGAGATGATAATTTTTTTTTTGAAGTAGAGATATGTTT 2600
QY	361 CCAACTCAGNAGCCTATGTGAACACTCAGCCAGCAAGCT -GGAAGTTTTTCTCTGTG 419
Db	2601 TTAATATTTAGAAGTTTATGTGAATTAATTAGTTAGTAAAGTTGGGAAGTTTTTTTTTG 2660
QY	420 ACCATGGGCTAATTGGTCTCCTCTCTCGATTTGGTCTTATCAGATAAAAACAAGTGAGT 479
Db	2661 ATTATGGGTTAATTGGTTTTTTTTTTTGGATTTGTTTTTATTTAGATAAAAAAAGTGGT 2720
QY	480 CATGCCACAGGATGCTATAGCCCATTTGATTTCTGGGATTTCTATGAGTGATGCTGATG 539
Db	2721 TAGTTATAGGATGTTTATAAGTTTATTGATTTTGGGATTTTATGAGTGATGTTGATG 2780
QY	540 ACTAAGCCAGGAGAGACTTATTTAAAGATCTCAGCATCTTTTACGCTGTTTAACTAGAGA 599
Db	2781 ATTAAGTAGGAGAGATTTATTTAAAGATTTTATTTAGTATTTTTTAGTTTGTTAATTTAGAGA 2840
QY	600 AAACCCGAGCATGATGATTAATAAGGGAATTTGAATGCGGTCCACCA 650



QY	85	AAATGCATTCTCTTCTACGGGATCTGGGAACCTTCCAAAGCTCCCTCCACAGTGGGAAT	144
Db	339	AAWYCMWKGKTTTMAAAWKGAMTKGAAARRGSMWYTCRSMWMMWTTWWAMMMWAW	398
QY	145	TTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGAGGCATCTTTCTTG	204
Db	399	TTWAWTTTMARKTYCMRGSTYKRMRGSCCYKGGGSMWYYWARSMARGRWGTSGMWW	458
QY	205	ACAGACAGAGTGGGAGCAGACAGCTGTCTACTTTCAGAGACTTCTTTCTGTATTC	264
Db	459	AAWTKGGYMAWYTGKKGAAARKTKGGKMYTTTKGYTMAWTTTWWATWMCMAWYY	518
QY	265	ATACCCCT	271
Db	519	TTYYVK	525
RESULT 11			
ACN37240		ACN37240 standard; DNA; 34796 BP.	
XX			
AC	ACN37240;		
DT	18-NOV-2004	(first entry)	
XX		Human periodontal disease related gene PLOD SEQ ID NO:150.	
DE		periodontal disease; polymorphism; ds; human; gene; SNP;	
KW		single nucleotide polymorphism.	
KW			
XX			
OS		Homo sapiens.	
XX			
FH	Key	Location/Qualifiers	
FT	misc_feature	869	
FT		/*tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide T,C"	
FT	misc_feature	1048	
FT		/*tag= b	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide T,C"	
FT	misc_feature	1103	
FT		/*tag= c	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide G,A"	
FT	misc_feature	5092	
FT		/*tag= d	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide A,G"	
FT	misc_feature	5333	
FT		/*tag= e	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide C,G"	
FT	misc_feature	5795	
FT		/*tag= f	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide T,C"	
FT	misc_feature	6328	
FT		/*tag= g	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide A,G"	
FT	misc_feature	6455	
FT		/*tag= h	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide A,G"	
FT	misc_feature	6489	
FT		/*tag= i	
FT		/standard_name= "Single nucleotide polymorphism"	
FT		/note= "Variable nucleotide C,G"	
FT	misc_feature	6616	
FT		/*tag= j	
FT		/standard_name= "Single nucleotide polymorphism"	
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			
FT			





XX 12-SEP-2003.  
XX  
XX 28-FEB-2003; 2003WO-US006235.  
XX  
XX 01-MAR-2002; 2002US-00087192.  
XX  
XX (SAGR-) SAGRES DISCOVERY.  
XX  
XX Morris DW;  
XX  
XX WPI; 2003-328604/31.  
XX  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
XX comprises a nucleotide sequence.  
XX  
XX Claim 1; SEQ ID NO 328; Opp; English.  
XX  
XX The present invention relates to novel DNA and protein sequences which  
XX are associated with carcinomas. The sequences are useful for: (i) for  
XX screening drug candidates; (ii) for screening of bioactive agent capable  
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
XX a bioactive agent capable of modulating the activity of CAP; (iv) for  
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
XX determining Carcinoma Associated (CA) gene copy number. In addition, the  
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of  
XX carcinoma including lymphoma. The present sequence is one such CA coding  
XX sequence. Note: This patent is an equivalent to basic patent  
XX US2002182586A1, for which no sequence data was published  
XX  
SQ Sequence 99886 BP; 25734 A; 17089 C; 18193 G; 27127 T; 0 U; 11743 Other;  
Query Match 5.4%; Score 34.8; DB 11; Length 99886;  
Best Local Similarity 54.8%; Pred. No. 18;  
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 101 ACGGGATCTGGGAACCTTCCAAAGCTGCCCTCCAGAGTGGGATTTCCACTCCTCTCT 160  
Db |||||  
26203 ACTTGAGTGGGAAGTTCAAGACCAAGCTGGCCCAACATGGCGAAACCCGTCTCATCTAA 26144  
QY 161 CAGCCGAGCACTGACCTCCAGCGGGGGAGGGCATCTTTTCTTGACAGAGCAGAGTGGG 220  
Db |||||  
26143 AAATACAGAAATTAGCGGGTGTGGTGGGGCCCTGTTCAGCTACTCAGGAGTCTG 26084  
QY 221 AGGCAG 226  
Db |||||  
26083 AGGCAG 26078

Search completed: August 4, 2005, 08:31:29  
Job time : 435.789 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 3067.89 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-7  
Perfect score: 650

Sequence: 1 ggcttggtgcgagatgtcc.....aattgaatgcgtccaccaa 650

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	42.8	6.6	834	BZ278233	CH230-343
C 2	42.4	6.5	845	CG856718	ZMMB8C026
C 3	42	6.5	673	AG053846	Pan trogl
C 4	40.6	6.2	764	CE213266	CE213266 tigr-gss-
C 5	39.2	6.0	454	CC111655	CC111655 ND.L.94D3.
C 6	38.8	6.0	306	AV215743	AV215743 AV215743
C 7	38.6	5.9	775	CO247511	AVENICOURT
C 8	37.8	5.8	569	CD217365	TgESTzy12
C 9	37.6	5.8	636	CE507575	tigr-gss-
C 10	37.6	5.8	1474	BF128374	601810241
C 11	37.4	5.8	637	CF895308	A0146D02-
C 12	37.4	5.8	640	CF895446	A0147H12-
C 13	37.4	5.8	644	AZ248574	RPCI-23-9
C 14	37.4	5.8	655	BU224166	603799875
C 15	37.4	5.8	725	BZ608804	WHABA547F
C 16	37.4	5.8	764	BQ203820	UI-R-DN1-
C 17	37.2	5.7	499	CK449936	893542 MA
C 18	37.2	5.7	556	CK453423	909135 MA
C 19	37.2	5.7	745	CK452363	907983 MA
C 20	37.2	5.7	935	BU124873	603149978
C 21	37	5.7	478	BQ210855	UI-R-D21-
C 22	37	5.7	532	BI278616	UI-R-CW0-
C 23	37	5.7	551	BQ199909	UI-R-DQ1-
C 24	37	5.7	569	CR314130	Medicago

25	36.6	5.6	331	4	BI298855	UI-R-CV2-
26	36.6	5.6	424	4	BG373719	UI-R-CV1-
27	36.6	5.6	559	6	CA505535	UI-R-FS1-
28	36.6	5.6	617	4	BM392304	UI-R-DQ1-
29	36.6	5.6	833	1	AL521194	AL521194
30	36.6	5.6	926	5	BX687092	BX687092
C 31	36	5.5	478	8	AQ222496	HS 3252_B
C 32	36	5.5	600	4	BG118574	602348317
C 33	35.8	5.5	555	2	BF389719	UI-R-BS2-
C 34	35.8	5.5	3367	3	AK047181	Mus muscu
C 35	35.6	5.5	228	2	BB002636	BB002636
C 36	35.6	5.5	491	2	BF155238	QV0-BF084
C 37	35.6	5.5	748	9	AG534765	Mus muscu
C 38	35.6	5.5	1504	4	BG396400	602459329
C 39	35.4	5.4	458	4	BI292863	UI-R-DOO-
C 40	35.4	5.4	459	2	BF440523	BS2900015
C 41	35.4	5.4	459	4	BM106126	BM106126
C 42	35.4	5.4	528	5	BP092719	BP092719
C 43	35.4	5.4	535	4	BI535053	398347 MA
C 44	35.4	5.4	570	2	AW289324	1511 Lewi
C 45	35.4	5.4	679	7	CK771529	959837 MA

## ALIGNMENTS

RESULT 1  
BZ278233/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BZ278233  
CH230-343L19.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone  
CH230-343L19, genomic survey sequence.  
BZ278233  
BZ278233.1 GI:24002829  
GSS.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 834)  
Zhao,S., Shetty,J., Shatsman,S., Tsagay,G., Geer,K.,  
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,  
Riggs,F., de Jong,P. and Fraser,C.M.  
Rat BAC End Sequences from Library CHORI-230 MboI segment  
Unpublished (1999)  
Other GSSs: CH230-343L19.TU  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Peter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering information.htm). BAC end  
page: http://www.tigr.org/cdb/bac\_ends/rat/bac\_end\_intro.html  
Plate: 343 row: L column: 19  
Seq primer: T7  
Class: BAC ends.

## FEATURES

Location/Qualifiers  
1. 834  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SeNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-343L19"  
/sex="female"  
/cell\_type="Brain"  
/clone\_lib="CHORI-230 Segment 2"  
/note="Vector: pTARBAC1.3; Site\_1: MboI; Site\_2: MboI;

Tue Aug 9 17:01:28 2005

us-09-899-276c-7.1st

```

CHORI-230 Rat (BN/SeNhsd/MCM) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match      6.6%; Score 42.8; DB 8; Length 834;
Best Local Similarity 54.4%; Pred. No. 0.41;
Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 457 TTATCGATATAAAACAAGTGAGTCTATGCCACAGAGTCTATTAAGCCCATTTGTTCTGGG 516
Db 523 TCATTATAGAAATATCTTTCATTCATGCCAGAAATATTACATAAGTGTAATTAATTGGGA 464

QY 517 ATTCTATGATGATGCTGATATGACTAAGCCAGGAGAGACTTATTATTAAGATCTCAGCAT 576
Db 463 AAAGGTTGATTTATTTATCAACACTATGAGATACAACTGGATTGAAGACATCAGCCT 404

QY 577 CTTTACGCTTTGTTAACTAGAGAAAACCCGAAGCATGA 614
Db 403 GTGTGTTCTCATTTAGTTAGCTAATAGTCTTACCTTGA 366

RESULT 2
CG856718/c
LOCUS      845 bp DNA linear GSS 19-NOV-2003
DEFINITION ZMMBBc0260N04f ZMMBBc (EcoRI) Zea mays genomic clone ZMMBBc0260N04
5', genomic survey sequence.
ACCESSION CG856718
VERSION    1
KEYWORDS  GSS.
SOURCE    GSS.
ORGANISM  Zea mays
           Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 845)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
Sequencing of the maize genome at PGIR (2003c)
Unpublished (2003)
Contact: Bharti.A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 445.
FEATURES
source
1..845
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBc0260N04"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc (EcoRI)"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN
Query Match      6.5%; Score 42.4; DB 9; Length 845;
Best Local Similarity 57.6%; Pred. No. 0.54;
Matches 76; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 427 GCTAATGTCCTCTCTCTGATGTCCTTATCAGATAAAACAAGTGAGTCATGCCA 486
Db 442 GATAATAGTTTATTTTGTCTTCATGATGTTGATACATATAACAAGTATTTATGCCA 383

QY 487 CAGGATGTCATAAGCCCATTAATCTGGATTTCTATGAGTATGCTGATGCTAAGC 546
Db 382 CATGAATATCCATGACCTTTAATGCTGGTAAACCATGTGACTTTGCTATATATAA 323

us-09-899-276c-7.1st

QY 547 CAGGACGAGACTT 558
Db 322 ACAGTGTGATTT 311

RESULT 3
AG053846/c
LOCUS      673 bp DNA linear GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-038B14.R, genomic survey sequence.
ACCESSION AG053846
VERSION    1
KEYWORDS  GSS.
SOURCE    GSS.
ORGANISM  Pan troglodytes (chimpanzee)
           Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 673)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..673
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-038B14.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      6.5%; Score 42; DB 9; Length 673;
Best Local Similarity 90.0%; Pred. No. 0.66;
Matches 45; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGCTTTGTGCCGAGATGTTCCAGCAGCAGCCCATGTGAGAGTCCCTGGC 50
Db 150 GGCTTTGTGCCGAAATGTTCCAGCAGCAGCCCATGTGAGAGTCTCGAGC 101

RESULT 4
CE213266/c
LOCUS      764 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-17000372782879 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION CE213266
VERSION    1
KEYWORDS  GSS.
SOURCE    GSS.
ORGANISM  Canis familiaris (dog)
           Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Fissipedia; Carnivora; Canidae; Canis.
1 (bases 1 to 764)

```

## AUTHORS

Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.

## TITLE

The dog genome: survey sequencing and comparative analysis

MEDLINE  
PUBMED

Science 301 (5641), 1898-1903 (2003)

## COMMENT

Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org  
Class: shotgun.

## FEATURES

source

Location/Qualifiers

1..764  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

## ORIGIN

Query Match 6.2%; Score 40.6; DB 9; Length 764;  
Best Local Similarity 81.6%; Pred. No. 1.8;  
Matches 71; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

## QY

3 CTTGTCGCGAGATGTTCCAGCAGCACGCCCATGTGAGAGCTCCCTGCTCCGCGGCCAGT 62

## Db

85 CTTGCACTGGGACATTCACGACGACGCCACCTGAGATGGGCC-GGCTCCAGTCCCGC 27

## QY

63 ATCTGGAATGCAGGCTCCAGCCAAATG 89

## Db

26 ATC-GGAATGCAGGCTCCAGCCAAACG 1

## RESULT 5

CC111655/c

## LOCUS

NDL.94D3.T7 Notre Dame Liverpool Aedes aegypti genomic clone

## DEFINITION

NDL.94D3, genomic survey sequence.

## ACCESSION

CC111655

## VERSION

CC111655.1 GI:29980710

## KEYWORDS

GSS.

## SOURCE

Aedes aegypti (yellow fever mosquito)

## ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;

Stegomyia.

1 (bases 1 to 454)

## REFERENCE

Loftus, B., Shetty, J., Knudson, D. and Severson, D.

BAC end sequencing of Aedes aegypti

Unpublished (2003)

Contact: Brendan Loftus

Department of Eukaryotic Genomics

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: enta@tigr.org

Library was provided by David Severson

Seq primer: T7

Class: BAC ends.

## FEATURES

source

Location/Qualifiers

1..454  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="Liverpool"  
/db\_xref="taxon:7159"  
/clone="NDL.94D3"

/clone\_lib="Notre Dame Liverpool"

/note="Vector: pECBAC1; Site 1: Hind III; The library was  
prepared from whole body tissue of newly hatched L1 larvae  
by David Severson at the University of Notre Dame and  
Hongbin Zhang"

## ORIGIN

Query Match 6.0%; Score 39.2; DB 8; Length 454;  
Best Local Similarity 50.5%; Pred. No. 3.9;  
Matches 95; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

## QY

436 TCTCCTTCTCTCGATTGTGGCTTATCAGATAAAAAACAAGTGAAGTCATGCCACAGGATGC 495

## Db

263 TTTTATTCTATGCGGGCTACATACACAAAAATAATTGGGTCCAGGCATTCAATAAA 204

## QY

496 TATAAGCCCATGATCTCTGGGATTCTATGAGTGATCTATGATGCTACTAGCCAGGAGAGA 555

## Db

203 AATCAAGGCTTTAATTCATGTATTGAAGATGAGGGGAGAAAAATAAAGGTTTGATCA 144

## QY

556 CTTATTAAAGATCTCAGCATCTTTTCAGCTTCTTAACCTAGAGAAAAACCGAAGCATGAC 615

## Db

143 CTTGTTCTTATCTGAGATATTGAATGATCTCTTTTAAAAATAACTAGATAATAATAG 84

## QY

616 TCGATTAT 623

## Db

83 TGGAAAAAT 76

## RESULT 6

AV215743

## LOCUS

AV215743 RIKEN full-length enriched, ES cells Mus musculus cDNA  
clone 2410152G06 3', similar to D13124 Rat mRNA for ATP synthase  
encoded by P2 gene, mRNA sequence.

## DEFINITION

AV215743

## ACCESSION

AV215743.1 GI:6156588

## VERSION

EST.

## KEYWORDS

Mus musculus

## SOURCE

Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 306)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T.,

Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F.,

Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,

Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusabe, M.,

Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,

Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,

Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, N.,

Suzuki, H., Suzuki, H., Takahashi, F., Tateno, M., Tomimaga, N.,

Tsunoda, Y., Watabiki, A., Watanabe, S., Yamamura, T., Yasunishi, A.,

Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al. 1999)

Unpublished (1999)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,

Matsuura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and

Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using RNA

polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,

Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDAM1163 row: a column: 22  
 High quality sequence start: 3  
 High quality sequence stop: 361.

## FEATURES

Location/Qualifiers

1. 775

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGS:30922965"

/tissue\_type="Chondrosarcoma Lung Metastasis cell lines"

/lab\_host="DH10B (T1 phage resistant)"

/clone.lib="NIH\_MGC\_212"

/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail sequence is: CAGTAAACCCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match	5.9%;	Score 38.6;	DB 7;	Length 775;
Best Local Similarity	55.2%;	Pred. No. 7.1;		
Matches	74;	Conservative	0;	Mismatches 60;
			Indels	Gaps
209	ACGAGAGTGGGAGGCACAGACGTGCTCACTTTCCAGAAAGACTTTCTTTTCTGATTCATAC	268		
509	AGCAGAAGGGGGAGTCCGTGTACCAACAACCTTTCTCACAACTTGGCTTTGTGTTTCAGCAC	450		
269	CCTTCACCTTCCCTCTGTTTACTGTCTCATATATGCAAGGCGCAAGTCACTTCCAGAGA	328		
449	CTTGGCGACGCCACGCTTTGNTGTGGACAGGTCCTCAGAGCCACTCTCTACCCAGGTA	390		
329	TGACAACCTCTCTCC	342		
389	GGAGCTCCCTCTTC	376		

RESULT	8	EST 20-MAY-2003
CD217365	569 bp	linear mRNA
CD217365	TgBSTz123ae02.v1	TGNas Tachyzoite cDNA Library
LOCUS	TgBSTz123ae02.v1	Toxoplasma gondii
DEFINITION	linear mRNA sequence.	

## ACCESSION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

TITLE	JOURNAL	COMMENT
1. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See also [1]
2. The Impact of Globalization on the Labor Market	Journal of International Economics	See also [2]
3. The Effect of Trade Liberalization on the Environment	Journal of Environmental Economics	See also [3]
4. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See also [4]
5. The Impact of Trade Liberalization on the Environment	Journal of Environmental Economics	See also [5]
6. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See also [6]
7. The Impact of Trade Liberalization on the Environment	Journal of Environmental Economics	See also [7]
8. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See also [8]
9. The Impact of Trade Liberalization on the Environment	Journal of Environmental Economics	See also [9]
10. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See also [10]

50

High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

## FEATURES

ity sequence: 589. 5. 775  
Location/Qualifiers  
1. 775  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="INAGS:30922965"  
/tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
/lab\_host="DHI10B (T1 phage resistant)"  
/clone\_lib="NIH MGC 212"  
/note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dt  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with Not I and then cloned  
directionally into pYX-Asc vector. The library tag  
direction located between the Not I site and the polyA tail  
is GATAGGCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

209	AGCAGAAGTGGGAGCGACACAGCTGTCACTTTCCAGAGACTTTCTTTCTGATTCATAC	266
509	AGCAGAAGGGGGAGTCCGTGACACAAACTTTCTCACAACTATGGCTTTGTGTCAGCAC	450
269	CCTTCACCTTCCCTGTGTTTACTGTCTGATATATGCAAGGCCAAGTCACTTTCCAGAGA	328
449	CCTGGCGACGCCACCGCTTTGTTGTGGACAGGTCCGAGGCCACTCTCTACCCAGGTA	390
329	TGACAACTCTCTCC	342
389	GGAGCTCCCTTTC	376

RESULT 7	775 bp	mRNA	linear	EST 23-JUN-2004
CO247511/c	AGSCNCTT 26525822	NIH MGC 212 Homo sapiens	CDNA clone	
LOCUS	CO247511	775 bp	mRNA	linear
DEFINITION	AGSCNCTT 26525822	NIH MGC 212 Homo sapiens	CDNA clone	

**SOURCE**  
**ORGANISM**

## REFERENCE AUTHORS

TITLE	JOURNAL	COMMENT

Seq primer: -40RP from Gibco.

## FEATURES

source  
 1. .569  
 /location="Toxoplasma gondii"  
 /mol\_type="mRNA"  
 /strain="Tachyzoite"  
 /db\_xref="taxon:5811"  
 /clone="TgESTy123e02.y1"  
 /dev\_stage="Tachyzoite"  
 /lab\_host="Electroten Blue cells (Stratagene)"  
 /clone\_lib="TgMAS Tachyzoite cDNA Library"  
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Kelian Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into Electroten Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."

ORIGIN

Query Match 5.8%; Score 37.8; DB 6; Length 569;  
 Best Local Similarity 60.0%; Pred. No. 11;  
 Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 246 AGACTTCTTTCTGATGATACCTTCCCTGCTGTTTACTGCTGATATGCA 305  
 Db |||||  
 321 AGTCTCTCTTTCTGTTTCAAACTCTTCACTTCTCGCTCCCTGCTCTTCTGAATG 380  
 QY 306 AAGCCGAGTCATCTTCCAGAGATGACACTCTCTCTGAAGTAG 350  
 Db |||||  
 381 GAGCGGAGACGGAAGGAAGCACACAAAGGACCCGGAAGAAG 425

RESULT 9  
 CE507575  
 LOCUS tigr-gss-dog-17000327371506 Dog Library Canis familiaris genomic,  
 DEFINITION genomic survey sequence.  
 ACCESSION CE507575  
 VERSION CE507575.1 GI:36824356  
 KEYWORDS GSS.  
 SOURCE Canis familiaris (dog)  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE 1 (bases 1 to 636)  
 AUTHORS Kirkness, E.F., Bafna V., Halpern, A.L., Levy, S., Remington, K.,  
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
 Venter, J.C.  
 TITLE The dog genome: survey sequencing and comparative analysis  
 JOURNAL Science 301 (5641), 1898-1903 (2003)  
 MEDLINE 22875432  
 PUBMED 14512627  
 COMMENT Contact: Kirkness EF  
 The Institute for Genomic Research  
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
 Rockville, MD 20850, USA  
 Tel: 301-838-0200  
 Fax: 301-838-0208  
 Email: ekirknes@tigr.org  
 Class: shotgun.  
 Location/Qualifiers  
 1. .636  
 /organism="Canis familiaris"  
 /mol\_type="genomic DNA"  
 /strain="Standard Poodle"  
 /db\_xref="taxon:9615"  
 /clone\_lib="Dog Library"

FEATURES  
 source

## ORIGIN

Query Match 5.8%; Score 37.6; DB 9; Length 636;  
 Best Local Similarity 49.5%; Pred. No. 13;  
 Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 444 TCTGGATTGCTTATCAGATAAACAAGTGAGTCATGCCACAGAGTCTCTATAAGCC 503  
 Db |||||  
 200 TCTGATGTTGTTAAATGGAATGATCCGAGCAAGGATAGATGTGATATGGA 259  
 QY 504 CATTGATTCGGGATCTTATGATGATGATGATGATGATGATGATGATGATGAT 563  
 Db CACAGAGTGGGCTGGTAGCATGAATGACAAGAGCAATCTCCAAGCAGGGCTGGTGT 319  
 QY 564 AGATCTCAGCATCTTTCAGCTTGTAACTAGAGAAAACCCGAGCATGCTGGATTAT 623  
 Db AACATGCCACTCTGCATCAGACTTTGGAACTCTAGTAATCTTTCATATGATGATGCTCA 379  
 QY 624 AAAGGGAATTCGAATG 639  
 Db 380 GAAAGAAATCCAATG 395

## RESULT 10

BF128374/c  
 LOCUS BF128374  
 DEFINITION mRNA sequence.  
 ACCESSION BF128374  
 VERSION BF128374.1 GI:10967414  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1474)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM891 row: b column: 12.

## FEATURES

source  
 1. .1474  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4053035"  
 /tissue\_type="leiomyosarcoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_46"  
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 5.8%; Score 37.6; DB 2; Length 1474;  
 Best Local Similarity 61.0%; Pred. No. 18;

Tue Aug 9 17:01:28 2005

Matches 61; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 189 GAGGCAATCTTTCTTGACAGAGAGAGTGGAGGACAGAGTGTCTACCTTTCCAGAAG 247

Db 1126 GGGCGTCTCTCTCTCGACACACATAGAGGGGGGCGCCCTTCTGTGCCCTACTCAGGAG 1067

QY 248 ACTTTCTTTTCTGATTACATACCTTCACCTTCCTCTGTGT 287

Db 1066 ACCTTCTTTCTGTTGGCTGTCTCTACTCTTGTGAT 1027

RESULT 11

CF895308

LOCUS

DEFINITION

A0146D02-5 NIA Mouse Undifferentiated ES Cell cDNA Library (Long 1)

Mus musculus cDNA clone NIA:A0146D02 IMAGE:30728005 5', mRNA

sequence.

CF895308

CF895308.1 GI:38162357

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 637)

Plao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL

MEDLINE

PUBMED

11544199

COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: A0146 row: D column: 02

Seq primer: M13 Reverse

High quality sequence stop: 637

POLYA=No.

Location/Qualifiers

1..637

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="129/Sv x 129/Sv-CP"

/db\_xref="niaEST:A0146D02-5"

/db\_xref="taxon:10090"

/clones="NIA:A0146D02 IMAGE:30728005"

/dev\_stage="R1 ES cells"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Undifferentiated ES Cell cDNA Library (Long 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site\_1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Kenneth R. Boheler (National Institute on Aging, USA). ES cells were cultured without feeder cells in the presence of LIF and BRL-conditioned media. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Loner-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes

ORIGIN

Query Match 5.8%; Score 37.4; DB 7; Length 637;

Best Local Similarity 54.8%; Pred. No. 15;

Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 225 AGACAGCTGTCACTTTCCAGAGACTTCTTTTCTGATTCATACCTTCACCTTCCTGT 284

Db 501 AGAGATTGTAATTAAGGATAATTCGCTTTCTGTTATTTCTCTTCTGTTCCATAC 560

QY 285 GTTTACTGTCTGATATATGCAAGGCCAGTCACTTTCCAGAGATGACAACTTCCTG 344

Db 561 AGTTCTTTGTTGTACATGCACCAACCAATCATTTACCAAGATGTCTCTTTATCAAT 620

QY 345 AAGTAGAGACATGCT 359

Db 621 AGTTACAGTCTGCT 635

RESULT 12

CF895446

LOCUS

DEFINITION

A0147H12-5 NIA Mouse Undifferentiated ES Cell cDNA Library (Long 1)

Mus musculus cDNA clone NIA:A0147H12 IMAGE:30728159 5', mRNA

sequence.

CF895446

CF895446.1 GI:38162495

EST.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 640)

Plao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.

Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method

Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL

MEDLINE

PUBMED

11544199

COMMENT

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: A0147 row: H column: 12

Seq primer: M13 Reverse

High quality sequence stop: 640

POLYA=No.

Location/Qualifiers

1..640

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="129/Sv x 129/Sv-CP"

/db\_xref="niaEST:A0147H12-5"

/db\_xref="taxon:10090"

/clones="NIA:A0147H12 IMAGE:30728159"

/dev\_stage="R1 ES cells"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Undifferentiated ES Cell cDNA Library (Long 1)"

/note="Vector: pCMV-SPORT6 (Invitrogen); Site\_1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Kenneth R. Boheler (National

FEATURES

source

Institute on Aging, USA). ES cells were cultured without feeder cells in the presence of LIF and BRL-conditioned media. Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen].  
 5'-P<sub>6</sub>CTAGTCTAGATCGAGCGCGCCCTTTTCTTTTCTTTT-3' from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to L<sub>1</sub>-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao."

```

ORIGIN
Query Match          5.8%; Score 37.4; DB 7; Length 640;
Best Local Similarity 54.8%; Pred. No. 15;
Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 225 AGACAGCTGTCACTTCCAGAGAACTTTCTTTCTGATTCATACACCTTCACCTTCCCTGT 284
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 501 AGAGAATTGTAATTGAAGGAATAATTCTGCTTTCTGTTATTCCTCTCGTTCATAC 560

QY 285 GTTACTGTCTGATATATGCAAGGCCAAGTCACTTCCAGAGATGACAACTCCCTTCCTG 344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 AGTTCTTGTGTTGCTACATCAACCAACAATCATCTACCAAGATGCTCTCTTTATCAAT 620

QY 345 AAGTAGAGACATGCT 359
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 AGTTACAGTCCTGGT 635

```

```

RESULT 13
LOCUS      AZ248574
DEFINITION RPCI-23-93C11.TJ RPCI-23 Mus musculus genomic clone RPCI-23-93C11,
            genomic survey sequence.
ACCESSION  AZ248574
VERSION     AZ248574.1 GI:8561693
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1 (bases 1 to 644)
AUTHORS     Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
            Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
            Jong, P. and Fraser, C.M.
            Mouse BAC End Sequences from Library RPCI-23
            Unpublished (1999)
            Other GSSs: RPCI-23-93C11.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@jeng.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
            Plate: 93 row: C column: 11
            Seq primer: SP6
            Class: BAC ends.
            Location/Qualifiers

```

## source

```

1. .644
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-93C11"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site: 1;
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies). "
```

```

ORIGIN
Query Match          5.8%; Score 37.4; DB 8; Length 644;
Best Local Similarity 54.8%; Pred. No. 15;
Matches 74; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 225 AGACAGCTGTCACTTCCAGAGAACTTTCTTTCTGATTCATACACCTTCACCTTCCCTGT 284
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 AGAGAATTGTAATTGAAGGAATAATTCTGCTTTCTGTTATTCCTCTCGTTCATAC 485

QY 285 GTTACTGTCTGATATATGCAAGGCCAAGTCACTTCCAGAGATGACAACTCCCTTCCTG 344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 AGTTCTTGTGTTGCTACATGACACCAACAATCATCTACCAAGATGCTCTCTTTATCAAT 545

QY 345 AAGTAGAGACATGCT 359
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 AGTTACAGTCCTGGT 560

RESULT 14
LOCUS      BU224166
DEFINITION 603799875F1 CSEQCHN23 Gallus gallus cDNA clone ChEST769f20 5', mRNA
            sequence.
ACCESSION  BU224166
VERSION     BU224166.1 GI:25458783
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
ORGANISM    Gallus gallus
REFERENCE   1 (bases 1 to 655)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
            Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
            A Comprehensive Collection of Chicken cDNAs
            Curr. Biol. 12 (22), 1965-1969 (2002)
            Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
            1. .655
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /strain="White Leghorn, Hisex"
            /db_xref="taxon:9031"
            /clone="ChEST769f20"
            /dev_stage="22"
            /lab_host="DH10B"
            /clone_lib="CSEQCHN23"

```

Tue Aug 9 17:01:28 2005

/note="Organ: heads; Vector: pBluescript II KS(+); Site\_1: EcoRI; Site\_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methyated C in the first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 5.8%; Score 37.4; DB 5; Length 655;  
Best Local Similarity 48.4%; Pred. No. 15;  
Matches 104; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 402 TGAAGTCTTCTCTGACCATGGCTAAATGGTCTCCTCTCTGGATTGTGGCTTATC 461  
Db 354 TAGAGCTTGTCTCTCCAGGAGAGATGATCTGCATTGTTGCTTGATATTCATATGC 413  
QY 462 AGATAAACAAGTGAAGTCAATGCCAAGGATGCTATTAAGCCCATGATCTCGGATTTCT 521  
Db 414 TGAAGAAGATGAAGATTTCTCGAGGTTCTCGGTATTACCGCTGTTTCAGTTGATCT 473  
QY 522 ATGAGTGATCTGATATGACTAAGCCAGGAGAGATTTTAAAGATCTCAGCATCTTTC 581  
Db 474 CCGCAGGCTGCTGCTAAGCCATCTCTGTGGAGATGACCTGGCAGCTTTGATGTTCC 533  
QY 582 AGCTTGTAACTAGAGAAACCCGAGCATGACT 616  
Db 534 TTTCAGTTATGACCTGAATCACTAAGGAACATT 568

RESULT 15  
BZ608804/c  
LOCUS  
DEFINITION  
BZ608804  
BZ608804.1 GI:31517365  
GSS.  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C.  
End-sequence profiling: Sequence-based analysis of aberrant genomes  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
22709111  
12788976  
Contact: Volik SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7065  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
http://www.genomex.com  
Class: BAC ends.

FEATURES  
source

Location/Qualifiers  
1..725  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-7111"

/sex="female"  
/clone\_lib="Human MCF7 breast cancer cell line library (MCF7\_1)"  
/note="Vector: pBAC1; Site\_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 5.8%; Score 37.4; DB 8; Length 725;  
Best Local Similarity 49.2%; Pred. No. 16;  
Matches 98; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 372 AAGCTATGTGAACACTCAGCCAGCAAGCTGGAAGTTTTTCTCTGTGACCATGGGCTAA 431  
Db 693 ATGCAGCTGTGGGAGATGCGACCTAAAGCCGATCTTCTGGAGGACAGCACTGGGCTGA 634  
QY 432 TTGGTCTCTCTCTGGATTGTGGCTTATCAGATATAAACAAGTGAAGTCAATGCCACAGGA 491  
Db 633 AAGGCTGAGTTGCTGCTTCTTCACTCTGAGAAATATAAATGAGCTTTGGGCTAAGAA 574  
QY 492 TGTCTATAAGCCCATTTGATTTCTGGATTCTATGAGTGATCTGATATGACTAAGCCAGGA 551  
Db 573 TGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 514  
QY 552 GAGACTTTATTTAAAGATCT 570  
Db 513 TGACATGATTTGATCTGT 495

Search completed: August 4, 2005, 14:32:44  
Job time : 3075.89 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 133.947 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-7  
Perfect score: 650  
Sequence: 1 ggcttgccgagatgttc.....aattgaatgcgtccaccaa 650

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.6	5.6	1497	4	US-09-220-132-94 Sequence 94, Appl
2	36.2	5.6	205044	4	US-09-949-016-15951 Sequence 15851, A
3	36.2	5.6	205044	4	US-09-949-016-15952 Sequence 15852, A
4	36.2	5.6	205044	4	US-09-949-016-15953 Sequence 15853, A
5	36.2	5.6	223471	4	US-09-949-016-12387 Sequence 12387, A
6	36.2	5.6	223471	4	US-09-949-016-12387 Sequence 12387, A
7	36.2	5.6	223471	4	US-09-949-016-12724 Sequence 12724, A
8	35.6	5.5	505	4	US-09-949-016-12725 Sequence 12725, A
9	33.8	5.2	141115	4	US-09-949-016-15639 Sequence 15639, A
10	33.6	5.2	760	4	US-09-586-106D-106 Sequence 106, App
11	33.4	5.1	1759	4	US-09-976-594-273 Sequence 273, App
12	33.4	5.1	1664976	4	US-08-916-421B-1 Sequence 1, Appl
13	33.4	5.1	1664976	4	US-09-692-570-1 Sequence 1, Appl
14	33.2	5.1	1144	4	US-09-640-211A-154 Sequence 154, App
15	33.2	5.1	121384	4	US-09-949-016-16944 Sequence 16944, A
16	33.2	5.1	392000	4	US-10-027-983-11 Sequence 11, Appl
17	33	5.1	1263	4	US-09-949-016-834 Sequence 834, App
18	33	5.1	1266	4	US-09-949-016-2294 Sequence 2294, App
19	33	5.1	312474	4	US-09-949-016-17434 Sequence 17434, App
20	32.8	5.0	120609	4	US-09-949-016-13915 Sequence 13915, A
21	32.6	5.0	7218	1	US-08-232-463-14 Sequence 14, Appl
22	32.4	5.0	601	4	US-09-949-016-160628 Sequence 160628, A
23	32.4	5.0	601	4	US-09-949-016-171536 Sequence 171536, A
24	32.4	5.0	1583	4	US-09-270-767-10265 Sequence 10265, A
25	32.4	5.0	1884	4	US-09-583-110-2476 Sequence 2476, App
26	32.4	5.0	1914	4	US-09-107-433-1788 Sequence 1788, App
27	32.4	5.0	151088	4	US-09-949-016-16240 Sequence 16240, A

C 28	32.4	5.0	298336	4	US-09-949-016-16600 Sequence 16600, A
C 29	32.4	5.0	385136	4	US-09-949-016-16073 Sequence 16073, A
C 30	32.2	5.0	832	4	US-09-621-976-2813 Sequence 2813, App
C 31	32.2	5.0	1200	4	US-08-956-171B-663 Sequence 663, App
C 32	32.2	5.0	1200	4	US-08-781-986A-663 Sequence 663, App
C 33	32.2	5.0	143644	4	US-09-949-016-15238 Sequence 15238, A
C 34	32.2	5.0	187848	4	US-09-949-016-12111 Sequence 12111, A
C 35	32	4.9	16404	4	US-09-949-016-16126 Sequence 16126, A
C 36	32	4.9	276237	4	US-09-949-016-17504 Sequence 17504, A
C 37	31.8	4.9	44789	4	US-09-949-016-13909 Sequence 13909, A
C 38	31.6	4.9	50186	4	US-09-949-016-14066 Sequence 14066, A
C 39	31.6	4.9	107800	4	US-09-949-016-13118 Sequence 13118, A
C 40	31.6	4.9	116425	4	US-09-949-016-11809 Sequence 11809, A
C 41	31.4	4.8	601	4	US-09-949-016-34983 Sequence 34983, A
C 42	31.4	4.8	601	4	US-09-949-016-154033 Sequence 154033, A
C 43	31.4	4.8	2118	4	US-09-690-454-16 Sequence 16, Appl
C 44	31.4	4.8	2955	4	US-09-248-796A-1987 Sequence 1987, App
C 45	31.4	4.8	30820	4	US-09-949-016-17145 Sequence 17145, A

ALIGNMENTS

RESULT 1

US-09-220-132-94  
; Sequence 94, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; PRIOR FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 94  
; LENGTH: 1497  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; LOCATION: (1)...(1497)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-220-132-94

Query Match  
Best Local Similarity 5.6%; Score 36.6; DB 4; Length 1497;  
Matches 48; Conservative 90; Mismatches 109; Indels 0; Gaps 0;

QY	25	ACAGCCCATGTGAGAGCTCCCTCGGCTCGGGCCCAAGTATCTGGAATCGAGCTCAGCC 84
DB	279	AAMCMCMWTKSMAAAMWMTTTTTTTTCMGSMRAARGGRARSYKRMWYMA 338
QY	85	AAATCATCTCTCTACGGATCTGGGAATCTCCAAAGCTCCCTCCTCAGATGGAT 144
DB	339	AAWYCMWKGKTTTMAAAAKGRAMYKGRRAARRGSMWYTCRSMWMTTMMWMAW 398
QY	145	TTCCACTCATTCTCTCAGCCAGCACTGACCTCCAGCGGGGAGGCGATCTTTTCTG 204
DB	399	TTWATTTTMAKTYTCMRGGSYTKRMGSCCYKGGGSMWYTYWARSWARGWTGSMW 458
QY	205	ACAGACGAGAAGTGGAGGCGAGACAGCTGTCTCACTTTCAGAGACTTTCTTCTGATC 264
DB	459	AAWTKGGYMAWYTTGKGAARRKTKGGKMYMTTTKGYTYNAWTTTTTWTWATMCCMAWY 518
QY	265	ATACCT 271
DB	519	TTTTYK 525

```

; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852

Query Match      5.6%; Score 36.2; DB 4; Length 205044;
Best Local Similarity 53.1%; Pred. No. 2.6;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTGGGAACTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 32809 CCATGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCACCTGCA 32750

QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGGAGGCACTTTTCTTGACAGAGCAGAAGTG 218
Db 32749 TATCTGCTGGCGTGAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 32690

QY 219 GGAGGCAGACAGCTGTCTCACTTTCCA 243
Db 32689 GACTTCACAAAGCTGCCCACTGCCA 32665

RESULT 4
US-09-949-016-15853/c
; Sequence 15853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15853
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15853

```

```

Query Match      5.6%; Score 36.2; DB 4; Length 205044;
Best Local Similarity 53.1%; Pred. No. 2.6;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTGGGAACTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 32809 CCATGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCACCTGCA 32750

QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGGAGGCACTTTTCTTGACAGAGCAGAAGTG 218
Db 32749 TATCTGCTGGCGTGAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 32690

QY 219 GGAGGCAGACAGCTGTCTCACTTTCCA 243
Db 32689 GACTTCACAAAGCTGCCCACTGCCA 32665

RESULT 5
US-09-949-016-12387/c
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

```

RESULT 2
US-09-949-016-15851/c
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match      5.6%; Score 36.2; DB 4; Length 205044;
Best Local Similarity 53.1%; Pred. No. 2.6;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTGGGAACTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 32809 CCATGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCACCTGCA 32750

QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGGAGGCACTTTTCTTGACAGAGCAGAAGTG 218
Db 32749 TATCTGCTGGCGTGAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 32690

QY 219 GGAGGCAGACAGCTGTCTCACTTTCCA 243
Db 32689 GACTTCACAAAGCTGCCCACTGCCA 32665

```

```

RESULT 3
US-09-949-016-15852/c
; Sequence 15852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15852
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```

;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
;; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
;; FILE REFERENCE: CL001307  
;; CURRENT APPLICATION NUMBER: US/09/949,016  
;; CURRENT FILING DATE: 2000-04-14  
;; PRIOR APPLICATION NUMBER: 60/241,755  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/237,768  
;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12387  
;; LENGTH: 223471  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(223471)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12387

Query Match 5.6%; Score 36.2; DB 4; Length 223471;  
Best Local Similarity 53.1%; Pred. No. 2.8;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 99 CTACGGGATCTGGGAACCTCCAAAGCTGCTCCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 3236 CCATGGGCGGGAGGTGGCAACATAGGTGCATGAGAGGGGACTGCCCCACCCCTGCA 3177  
  
QY 159 CTCAGCCAGCAGCTGACCTCCAGCGGGGAGGCGATCTTTTCTTGACAGACGAGAAGTG 218  
Db 3176 TATCTGCTGGCGCTGAAATTTCTCAGGACAGCTAGCCCTGCTTTCAGAGATGAGAG 3117  
  
QY 219 GGAGGCAGACAGCTGTCACCTTTCCA 243  
Db 3116 GACTTCACAAAGCTGCCCACTGCCA 3092

RESULT 6  
US-09-949-016-12724/c  
; Sequence 12724, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12724  
; LENGTH: 223471  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(223471)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12724

Query Match 5.6%; Score 36.2; DB 4; Length 223471;  
Best Local Similarity 53.1%; Pred. No. 2.8;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTGGGAACCTCCAAAGCTGCTCCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 3236 CCATGGGCGGGAGGTGGCAACATAGGTGCATGAGAGGGGACTGCCCCACCCCTGCA 3177  
  
QY 159 CTCAGCCAGCAGCTGACCTCCAGCGGGGAGGCGATCTTTTCTTGACAGACGAGAAGTG 218  
Db 3176 TATCTGCTGGCGCTGAAATTTCTCAGGACAGCTAGCCCTGCTTTCAGAGATGAGAG 3117  
  
QY 219 GGAGGCAGACAGCTGTCACCTTTCCA 243  
Db 3116 GACTTCACAAAGCTGCCCACTGCCA 3092

RESULT 7  
US-09-949-016-12725/c  
; Sequence 12725, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12725  
; LENGTH: 223471  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(223471)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12725

Query Match 5.6%; Score 36.2; DB 4; Length 223471;  
Best Local Similarity 53.1%; Pred. No. 2.8;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 99 CTACGGGATCTGGGAACCTCCAAAGCTGCTCCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 3236 CCATGGGCGGGAGGTGGCAACATAGGTGCATGAGAGGGGACTGCCCCACCCCTGCA 3177  
  
QY 159 CTCAGCCAGCAGCTGACCTCCAGCGGGGAGGCGATCTTTTCTTGACAGACGAGAAGTG 218  
Db 3176 TATCTGCTGGCGCTGAAATTTCTCAGGACAGCTAGCCCTGCTTTCAGAGATGAGAG 3117  
  
QY 219 GGAGGCAGACAGCTGTCACCTTTCCA 243  
Db 3116 GACTTCACAAAGCTGCCCACTGCCA 3092

RESULT 8  
US-09-621-976-15639  
; Sequence 15639, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm

```

; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match          5.5%; Score 35.6; DB 4; Length 505;
Best Local Similarity 14.3%; Pred. No. 0.096;
Matches 65; Conservative 181; Mismatches 205; Indels 5; Gaps 1;

QY 24 CACAGCCCATGTCAGAGCTCCTCGCTCCGGCCCGACATCTGGAATGCAGGCTCCAGC 83
Db 2 MAYMSRYRTSSKSYGKYSRWSMKYSGRKYTSYANGRGKYGTSGRGCGSGRWSKRGY 61
QY 84 CAAATGCAATCTCTCTACGGAGTCTGGGAATTCCTCAAGATGCTCCCTCAGAGTGGAA 143
Db 62 RYSYWGKWSKWSKWKYSGRKGTSSKTRKYRTYTSKCRKTTCKYRGWSKWKRMKRR 121
QY 144 TTTCACACTACTCTCTCAGCCAGCAGCTGACTCCAGCGGGGAGGCGCATCTTTCTT 203
Db 122 KMYRYRMKYSYASYSYRRCKRYTGMRGWYCKRCKMSSTRYMTRYTRYWMTGAC 181
QY 204 GACAGCACAAGTGGG-AGGCAGACAGCTGTCATCTTCCAGAAGACTTTCTTTC 258
Db 182 YGSGMCKGSRKYSYGRWYKGYKTMCTSKYKSSKMSYKSSMCYTYMYCYTYWCT 241
QY 259 TGATTCATACCTTCACTTCCCTCGTGTGTACTGTCTGATATATGAAAGCCAGTCAC 318
Db 242 CTYKKSYYTCKRSYYTSTSKGWTGKSRYSYTWMSKSYTWGSKWKWMSAGAW 301
QY 319 TTTCAGAGATGACAATCTCTCTGAGCTAGAGACATGCTTCCAACTCAGAGACCTA 378
Db 302 YAMMSWCMARMCMAGMRSAWKKCRKYMIMAKSCMYCAKMSCSARSASRSCCTT 361
QY 379 TGTGAACAACCTCAGCAGCAAGCTGGAAGTGTCTCTGTGACCATGGCTAATGTGCT 438
Db 362 TKYKMTTKCYCKYRCWMSWCSWCAICYTGYWCYVYTKTWGTCMTKAACTYGTYSCT 421
QY 439 CTTCTCTGGATTTGGCTTATCATGATAAAACAAG 474
Db 422 KMYMTTCTCTWYGTTRATTTGAGRRGTARRAG 457

RESULT 9
US-09-949-016-17490
; Sequence 17490, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17490
; LENGTH: 14115
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(14115)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17490

; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match          5.2%; Score 33.6; DB 4; Length 760;
Best Local Similarity 52.9%; Pred. No. 0.62; Indels 0; Gaps 0
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0

QY 196 CTTTCTTTCAGACGACGAGTGGGAGGCGACAGCTGTCACCTTTCCAGAGACTTTCTT 255
Db 285 CTATTCAGGATACACCAATTTGTTGTCACGAGGACGAGGAGAACTACATTCAC 344
QY 256 TTCTGATTCATACCTTCCCTCCCTGTTTACTGCTGTATATATCAAGGCGCAAGT 315
Db 345 TTGTCTGTATGGGACATATGCTTTCAAGTGAAGTGGCTGCTGCTGCTGCTGCTGCT 404
QY 316 CACTTTCACAGATGA 331
Db 405 CACCTTCCAAAGATGA 420

RESULT 11
US-09-976-594-273
; Sequence 273, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12

```

```
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 273
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 404601.11
; NAME/KEY: unsure
; LOCATION: 1731..1736, 1739, 1742-1743, 1747, 1749, 1751-1752, 1754, 1756
; OTHER INFORMATION: a, t, c, g, or other
;
US-09-976-594-273

Query Match          5.1%; Score 33.4; DB 4; Length 1759;
Best Local Similarity 55.7%; Pred. No. 1.2;
Matches 64; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 146 TCCTCTCCTCTCTCTCAGCCAGACGTGACCTCCCGGGGAGGCGCATCTTTTCTTGA 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 TCCTCTCTCTCTCTCTTTTCTCCCTCCCTGATAGCGAAGAGTGTTTCTTTACCC 1334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 206 CAGACGAGAGTGGGAGGAGAGAGCTGCTCACTTCCAGAGACTTCTTTCTG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 CTTGACAGAGTGGTCTCTGCCACCTTATATTCCCTGTAGTCCCTCTTCTG 1389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

```
Query Match 5.1%; Score 33.4; DB 4; Length 1664976;
Best Local Similarity 57.0%; Pred. No. 91;
Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 262 TTCATACCTTCACCTCCCTGCTTACTGTCGATATATGCAAGGCCAAGTCACCTTT 321
DB 263285 TTCATAGCTTTAACTTTTCCCTGTTTACTGCCCTTAAACAGCATCAATTAACATTTT 263226

QY 322 CCAGAGATGACACTCCTTCCTGAGTAGACATGCTTCCAACT 368
DB 263225 TCAGAAATGATATCCCAACTGACACTTGGCCCTTCTTCTCACT 263179
```

## RESULT 13

```
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
```

```
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (582442)..(582442)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
```

Query Match 5.1%; Score 33.4; DB 4; Length 1664976;

Best Local Similarity 57.0%; Pred. No. 91;

Matches 61; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```
QY 262 TTCATACCCCTTCACCTTCCTGTTTACTGTCTGATATATCAAGGCCAAGTCACATTT 321
Db 263285 TTCATAGCTTTAACTTTCTGTTTACTGCTCTAAACAGCATCAATTAACATTTT 263226
QY 322 CCAGAGATGACAACTTCCTTCCTGGAAGTAGAGACATGCTTCCAACT 368
Db 263225 TCAGAAATGCATCCCAACCTGAACACTTGGCCTTTCTTCTCACT 263179
```

RESULT 14  
US-09-640-211A-154  
; Sequence 154, Application US/09640211A

Tue Aug 9 17:01:28 2005

us-09-899-276c-7.rni

Qy 614 ACTGGATTATTAAGGGAATTTGAATG 639  
 Db 114410 CCAACATTAGAAACTGAATTTCAATG 114435  
 Search completed: August 4, 2005, 14:41:07  
 Job time : 142.947 secs

Patent No. 6833446  
 GENERAL INFORMATION:  
 APPLICANT: Wood, Marion  
 APPLICANT: Shenk, Michael A.  
 APPLICANT: McGrath, Annette  
 APPLICANT: Glenn, Matthew  
 TITLE OF INVENTION: Compositions and Methods for the  
 TITLE OF INVENTION: Modification of Gene Transcription  
 FILE REFERENCE: 11000.1021C1U  
 CURRENT APPLICATION NUMBER: US/09/640,211A  
 CURRENT FILING DATE: 2000-08-16  
 NUMBER OF SEQ ID NOS: 2368  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 154  
 LENGTH: 1144  
 TYPE: DNA  
 ORGANISM: Eucalyptus grandis  
 US-09-640-211A-154

Query Match 5.1%; Score 33.2; DB 4; Length 1144;  
 Best Local Similarity 55.1%; Pred. No. 1.1;  
 Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 66 TGAATGCGAGCTCCAGCCAAATGCAATCTCTTCTACGGGATCTGGGAATTTCCAAAGCT 125  
 Db 573 TCGAGTGCTGGCTCCACCGGAGCGATACCGGACGCGAGCGTGCAAGGACGGGCAAGCT 632

Qy 126 GCCTCTCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCGCAGCACTGACCTCCCGCAGC 183  
 Db 633 GCCGCGCGCGTCTGCTTCTTCCGCACTTCCCGGACCGAGCTCCGGGTCTCTCCCGC 690

RESULT 15  
 US-09-949-016-16944  
 Sequence 16944 Application US/09949016  
 Patent No. 6812339  
 GENERAL INFORMATION:  
 APPLICANT: VENTER, J. Craig et al.  
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: CL001307  
 CURRENT APPLICATION NUMBER: US/09/949,016  
 CURRENT FILING DATE: 2000-04-14  
 PRIOR APPLICATION NUMBER: 60/241,755  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/237,768  
 PRIOR FILING DATE: 2000-10-03  
 PRIOR APPLICATION NUMBER: 60/231,498  
 PRIOR FILING DATE: 2000-09-08  
 NUMBER OF SEQ ID NOS: 207012  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 16944  
 LENGTH: 121384  
 TYPE: DNA  
 ORGANISM: Human  
 US-09-949-016-16944

Query Match 5.1%; Score 33.2; DB 4; Length 121384;  
 Best Local Similarity 47.6%; Pred. No. 21;  
 Matches 98; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 434 GGTCTCTCTCTCTGATTTGCTTATCAGATAAAACAAAGTGAGTGATGCCACAGGATG 493  
 Db 114230 GCTCACCTCTGGTGTGTGTGACCATCCCNAAACCAAGGGCGAGAGCAGAGATG 114289

Qy 494 TCTATAAGCCCATTTGATTTGGGATTTCTATGAGTGATGCTGATATGACTAAGCCAGGAGA 553  
 Db 114290 TCCCTGTGAGGATTTGATGATATACTGATATAGTTTGTGTTTAAATGAAAAACAAGAGC 114349

Qy 554 GACCTATTAAAGATCTCAGCATCTTTCAGCTTGTAAACCTAGAGAAACCCGAAGCATG 613  
 Db 114350 AAACAAAATTTTAAATCCATGATTTAGCCCATTCATTTCTAGACCTAATTTATTGATC 114409

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 1159.47 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-7  
Perfect score: 650  
Sequence: 1 ggcttgccgagatttcc.....aattgaatcggtccaccaa 650

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues  
Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	650	100.0	650	9	US-09-899-276-7
2	616.6	94.9	11793	19	US-10-685-705-4
3	608.6	93.6	3221	19	US-10-685-705-3
4	401.8	61.8	5926	15	US-10-311-455-1626
5	365.4	56.2	5926	15	US-10-311-455-1625
6	38.2	5.9	6557	15	US-10-311-455-973
7	36.6	5.6	1497	21	US-10-831-704-94

Sequence 254387,	US-10-027-632-254387	13	1266	5.6	36.2	8
Sequence 254388,	US-10-027-632-254388	13	1266	5.6	36.2	9
Sequence 254389,	US-10-027-632-254389	13	1266	5.6	36.2	10
Sequence 254390,	US-10-027-632-254390	13	1266	5.6	36.2	11
Sequence 254387,	US-10-027-632-254387	17	1266	5.6	36.2	12
Sequence 254388,	US-10-027-632-254388	17	1266	5.6	36.2	13
Sequence 254389,	US-10-027-632-254389	17	1266	5.6	36.2	14
Sequence 254390,	US-10-027-632-254390	17	1266	5.6	36.2	15
Sequence 529, App	US-10-087-192-529	13	211257	5.5	35.8	16
Sequence 264907,	US-10-027-632-264907	13	1319	5.4	35.4	17
Sequence 264907,	US-10-027-632-264907	17	1319	5.4	35.4	18
Sequence 105632,	US-10-027-632-105632	13	671	5.4	35.2	19
Sequence 105632,	US-10-027-632-105632	17	671	5.4	35.2	20
Sequence 17506, A	US-10-027-632-17506	13	740	5.4	35.2	21
Sequence 87, Appl	US-10-027-632-17506	17	740	5.4	35.2	22
Sequence 16668, A	US-10-175-523-87	14	130427	5.4	35	23
Sequence 150617,	US-10-027-632-16668	13	689	5.4	34.8	24
Sequence 150617,	US-10-027-632-16668	17	689	5.4	34.8	25
Sequence 328, App	US-10-027-632-150617	13	926	5.4	34.8	26
Sequence 95, Appl	US-10-087-192-328	17	926	5.4	34.8	27
Sequence 14649, A	US-10-461-862-95	21	192992	5.4	34.8	28
Sequence 219, App	US-10-787-701-14649	19	413	5.3	34.6	29
Sequence 465, App	US-10-221-613-219	17	19576	5.3	34.6	30
Sequence 984, App	US-09-938-842A-465	9	1947	5.3	34.4	31
Sequence 301, App	US-10-311-455-984	15	1947	5.3	34.4	32
Sequence 474, App	US-10-108-260A-301	17	2243	5.2	34	33
Sequence 711, App	US-09-876-143-474	11	502	5.2	33.8	34
Sequence 89, Appl	US-10-311-455-711	15	6759	5.2	33.8	35
Sequence 416, App	US-10-221-613-89	17	15923	5.2	33.8	36
Sequence 106, App	US-10-221-714A-416	18	24259	5.2	33.8	37
Sequence 106, App	US-10-395-607-106	17	760	5.2	33.6	38
Sequence 1023, Ap	US-10-799-870-106	19	760	5.2	33.6	39
Sequence 1294, Ap	US-09-938-842A-1023	9	885	5.2	33.6	40
Sequence 191, App	US-09-938-842A-1294	11	885	5.2	33.6	41
	US-10-311-455-1294	15	16509	5.2	33.6	42
	US-10-388-934-191	17	14279	5.1	33.4	43
						44
						45

ALIGNMENTS

RESULT 1  
US-09-899-276-7  
; Sequence 7, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Delius, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899, 276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 650  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-7

Query Match 100.0%; Score 650; DB 9; Length 650;  
Best Local Similarity 100.0%; Pred. No. 1e-209;  
Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCTTGTGCGAGATGTTCCAGCA	CAGCCCAATGTGAGAGCTCCCTGGCTCCGGGCCCA	60
Db	1	GGCTTGTGCGAGATGTTCCAGCA	CAGCCCAATGTGAGAGCTCCCTGGCTCCGGGCCCA	60
QY	61	GTATCTGGAAATGAGGCTCCAGCA	AAATGCAATCTCTTACGGGATCTGGAACTTCCA	120
Db	61	GTATCTGGAAATGAGGCTCCAGCA	AAATGCAATCTCTTACGGGATCTGGAACTTCCA	120
QY	121	AAGCTGCTCTCCTCAGAGTGGAA	TTTCCACTCCTCAGCCAGCAGCTGCTCTCC	180
Db	121	AAGCTGCTCTCCTCAGAGTGGAA	TTTCCACTCCTCAGCCAGCAGCTGCTCTCC	180
QY	181	AGCGGGGAGGGCACTTTCTTGA	CAGAGCAAGTGGAGGAGCAGACAGCTGTT	240
Db	181	AGCGGGGAGGGCACTTTCTTGA	CAGAGCAAGTGGAGGAGCAGACAGCTGTT	240
QY	241	CCAGAAAGCTTTCTTTCTGATT	CAACCTTCCCTTCCCTGTTTACTGCTGATAT	300
Db	241	CCAGAAAGCTTTCTTTCTGATT	CAACCTTCCCTTCCCTGTTTACTGCTGATAT	300
QY	301	ATGCAAGGCAAGTCACTTTCCAG	AGATGCAACTCTCTGAGTAGACATGCTT	360
Db	301	ATGCAAGGCAAGTCACTTTCCAG	AGATGCAACTCTCTGAGTAGACATGCTT	360
QY	361	CCAACACTCAGAGCTTATGTGA	CACTCAGCCAGCAAGCTGGGAAGTCTCTG	420
Db	361	CCAACACTCAGAGCTTATGTGA	CACTCAGCCAGCAAGCTGGGAAGTCTCTG	420
QY	421	CCATGGCTAAATGGTCTCTCT	CTGATTTGAGTGTGATTAAGAAAGTGA	480
Db	421	CCATGGCTAAATGGTCTCTCT	CTGATTTGAGTGTGATTAAGAAAGTGA	480
QY	481	ATGCCAGAGATGTTATAGCCAT	TGATTTCTGGGATTTCTAGTGTGATGTA	540
Db	481	ATGCCAGAGATGTTATAGCCAT	TGATTTCTGGGATTTCTAGTGTGATGTA	540
QY	541	CTAAGCCAGAGAGACTTATTA	AGATCTCAGCATCTTTTAAACCTAGAGAA	600
Db	541	CTAAGCCAGAGAGACTTATTA	AGATCTCAGCATCTTTTAAACCTAGAGAA	600
QY	601	AACCCGAAGCATGATGATTA	AAAGGAAATGAATGCGTCCACCAA	650
Db	601	AACCCGAAGCATGATGATTA	AAAGGAAATGAATGCGTCCACCAA	650
RESULT 2				
US-10-685-705-4				
; Sequence 4, Application US/10685705				
; Publication No. US20040177387A1				
; GENERAL INFORMATION:				
; APPLICANT: University of Kentucky Research Foundation				
; APPLICANT: JAYAKRISHNA, Ambati				
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular				
; FILE REFERENCE: 050229-0415				
; CURRENT APPLICATION NUMBER: US/10/685,705				
; PRIOR FILING DATE: 2003-10-16				
; PRIOR APPLICATION NUMBER: 60/422,096				
; PRIOR FILING DATE: 2002-10-30				
; NUMBER OF SEQ ID NOS: 10				
; SOFTWARE: Patentin version 3.2				
; SEQ ID NO 4				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-685-705-4				
Query Match 94.9%; Score 616.6; DB 19; Length 11793;				
Best Local Similarity 97.7%; Pred. No. 1.2e-197;				
Matches 636; Conservative 0; Mismatches 14; Indels 1; Gaps 1;				
QY	1	GGCTTGTGCGAGATGTTCCAGCA	CAGCCCAATGTGAGAGCTCCCTGGCTCCGGGCCCA	60

Db	4797	GGCTTGTGCGAGATGTTCCAGCA	CAGCCCAATGTGAGAGCTCCCTGGCTCCGGGCCCA	4856
QY	61	GTATCTGGAAATGAGGCTCCAGCA	AAATGCAATCTCTTACGGGATCTGGAACTTCCA	120
Db	4857	GTATCTGGAAATGAGGCTCCAGCA	AAATGCAATCTCTTACGGGATCTGGAACTTCCA	4916
QY	121	AAGCTGCTCTCCTCAGAGTGGAA	TTTCCACTCCTCAGCCAGCAGCTGCTCTCC	180
Db	4917	AAGCTGCTCTCCTCAGAGTGGAA	TTTCCACTCCTCAGCCAGCAGCTGCTCTCC	4976
QY	181	AGCGGGGAGGGCACTTTCTTGA	CAGAGCAAGTGGAGGAGCAGACAGCTGTT	240
Db	4977	AGCGGGGAGGGCACTTTCTTGA	CAGAGCAAGTGGAGGAGCAGACAGCTGTT	5036
QY	241	CCAGAAAGCTTTCTTTCTGATT	CAACCTTCCCTTCCCTGTTTACTGCTGATAT	300
Db	5037	CCAGAAAGCTTTCTTTCTGATT	CAACCTTCCCTTCCCTGTTTACTGCTGATAT	5096
QY	301	ATGCAAGGCAAGTCACTTTCCAG	AGATGCAACTCTCTGAGTAGACATGCTT	360
Db	5097	ATGCAAGGCAAGTCACTTTCCAG	AGATGCAACTCTCTGAGTAGACATGCTT	5156
QY	361	CCAACACTCAGAGCTTATGTGA	CACTCAGCCAGCAAGCTGGGAAGTCTCTG	419
Db	5157	CCAACACTCAGAGCTTATGTGA	CACTCAGCCAGCAAGCTGGGAAGTCTCTG	5216
QY	420	ACCATGGCTAAATGGTCTCTCT	CTGATTTGAGTGTGATTAAGAAAGTGA	479
Db	5217	ACCATGGCTAAATGGTCTCTCT	CTGATTTGAGTGTGATTAAGAAAGTGA	5276
QY	480	CATGCCAGAGATGTTATAGCCAT	TGATTTCTGGGATTTCTAGTGTGATGTA	539
Db	5277	CATGCCAGAGATGTTATAGCCAT	TGATTTCTGGGATTTCTAGTGTGATGTA	5336
QY	540	ACTAAGCCAGAGAGACTTATTA	AGATCTCAGCATCTTTTAAACCTAGAGA	599
Db	5337	ACTAAGCCAGAGAGACTTATTA	AGATCTCAGCATCTTTTAAACCTAGAGA	5396
QY	600	AAACCCGAGCATGATGATTA	AAAGGAAATGAATGCGTCCACCAA	650
Db	5397	AAACCCGAGCATGATGATTA	AAAGGAAATGAATGCGTCCACCAA	5447
RESULT 3				
US-10-685-705-3				
; Sequence 3, Application US/10685705				
; Publication No. US20040177387A1				
; GENERAL INFORMATION:				
; APPLICANT: University of Kentucky Research Foundation				
; APPLICANT: JAYAKRISHNA, Ambati				
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular				
; FILE REFERENCE: 050229-0415				
; CURRENT APPLICATION NUMBER: US/10/685,705				
; PRIOR FILING DATE: 2003-10-16				
; PRIOR APPLICATION NUMBER: 60/422,096				
; PRIOR FILING DATE: 2002-10-30				
; NUMBER OF SEQ ID NOS: 10				
; SOFTWARE: Patentin version 3.2				
; SEQ ID NO 3				
; LENGTH: 3221				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-685-705-3				
Query Match 93.6%; Score 608.6; DB 19; Length 3221;				
Best Local Similarity 97.7%; Pred. No. 3e-195;				
Matches 628; Conservative 0; Mismatches 14; Indels 1; Gaps 1;				
QY	9	CCGAGATGTTTCCAGCAGACAG	CCCCCATGTGAGAGCTCCCTGGCTCCGGGCCCA	68
Db	1	CCGAGATGTTTCCAGCAGACAG	CCCCCATGTGAGAGCTCCCTGGCTCCGGGCCCA	60

QY 69 AATGAGGCTCCAGCCAAATGCAATCTCTTACGGGATCTGGGAATCTTCCAAAGCTGCC 128  
DB 61 AATGAGGCTCCAGCCAAATGCAATCTCTTACGGGATCTGGGAATCTTCCAAAGCTGCC 120  
QY 129 TCCTCAGAGTGGGAATTTCCACTCACTTCTCTCA CGCCAGCACTGACCTCCAGCGGGG 188  
DB 121 TCCTCAGAGTGGGAATTTCCACTCACTTCTCTCA CGCCAGCACTGACCTCCAGCGGGG 180  
QY 189 AGGGCACTCTTTCTTCACAGAGCAGAGTGGGAGGAGACAGCTGTCATTTTCCAGAAGA 248  
DB 181 AGGGCACTCTTTCTTCACAGAGCAGAGTGGGAGGAGACAGCTGTCATTTTCCAGAAGA 240  
QY 249 CTTTCTTTTCTGATTCATACCTTCCACTTCCCTGCTGTTTACTGCTGATATATGCAAG 308  
DB 241 CTTTCTTTTCTGATTCATACCTTCCACTTCCCTGCTGTTTACTGCTGATATATGCAAG 300  
QY 309 GCCAAGTCACTTTCCAGAGATGACAACTCTCTTCTGAAAGTAGAGACATGCTTCCAACT 368  
DB 301 GCCAAGTCACTTTCCAGAGATGACAACTCTCTTCTGAAAGTAGAGACATGCTTCCAACT 360  
QY 369 CAGAAGCTATGTGAACACTCAGCCAGCAAGCT -GGAAGTCTTCTGTCGACCATGG 427  
DB 361 CAGAAGCTATGTGAACACTCAGCCAGCAAGCTGGAAGTCTTCTGTCGACCATGG 420  
QY 428 CTAATTGGTCTCTCTCTGGATTTGGGCTTATCAGATAAAACAAGTGAGTCATGCCAC 487  
DB 421 CTAATTGGTCTCTCTCTGGATTTGGGCTTATCAGATAAAACAAGTGAGTCATGCCAC 480  
QY 488 AGAATGCTATAAGCCATGATTTGGGATTTCTGAGTATGCTGATATGACATAAGCC 547  
DB 481 AGAATGCTATAAGCCATGATTTGGGATTTCTGAGTATGCTGATATGACATAAGCC 540  
QY 548 AGGAGAGCTTATTTAAAGATCTCAGCATCTTTCAGCTTTTAACTAGAGAAACCCGA 607  
DB 541 AGGAGAGCTTATTTAAAGATCTCAGCATCTTTCAGCTTTTAACTAGAGAAACCCGA 600  
QY 608 AGCATGACTGGATTATAAGGGAATTTGAATGGGTCCACCAA 650  
DB 601 AGCATGACTGGATTATAAGGGAATTTGAATGGGTCCACCAA 643

## RESULT 4

US-10-311-455-1626/c

; Sequence 1626, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1626

; LENGTH: 5926

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1626

Query Match

Best Local Similarity 61.8%; Score 401.8; DB 15; Length 5926;

Matches 501; Conservative 0; Mismatches 147; Indels 1; Gaps 1;

QY 3 CTTGTGCGAGATGTTCCAGCAGACCCCATGTGAGAGTCCCTGGCTCCGGGCCAGT 62  
DB 3684 CTTATACCGAATATTTCCACACACACCCCATATATAAACTCCCTAACTCCGACCAAT 3625  
QY 63 ATCTGGAATGCAAGTCCAGCCAAATGCAATCTCTTCTACGGGATCTGGGAATCTTCCAAA 122  
DB 3624 ATCTGGAATGCAAGTCCAGCCAAATGCAATCTCTTCTACGGGATCTGGGAATCTTCCAAA 3565  
QY 123 GCTGCTCTCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCAGCAGCTGACCTCCAG 182  
DB 3564 ACTACTCTCTCAAAATAAAATTTCCACTCACTTCTCTCAGCCAGCAGCTGACCTCCAG 3505  
QY 183 CGGGGAGGGGCACTTTTCTTGACAGAGCAGAAAGTGGGAGGAGCAGAGCTGTCACTTTCC 242  
DB 3504 CGAAAAAACAATCTTTTCTTAACAACAACAATAAATAAACAACAATATCACTTTCC 3445  
QY 243 AGAAGACTTTCTTTCTGATTCATACCTTCACTTCTCTGCTGTTTACTGTCTGATATAT 302  
DB 3444 AAAAATCTTTCTTTCTAATTCATACCTTCACTTCTCTGCTGTTTACTGTCTGATATAT 3385  
QY 303 GCAGAGGCGAAGTCACTTTCCAGAGATGACAACTCTTCTGAAAGTAGAGATGCTTCC 362  
DB 3384 ACAAAAACCAATCACTTTTCCAAAAATAACAACCTCTTCTTAAAAATAAACAATACTTCC 3325  
QY 363 AACACTCAGAGAGCTATGTGAACACTCAGCCAGCAAGCT -GGAAGTCTTCTCTGTGAC 421  
DB 3324 AACACTCAGAGAGCTATGTGAACACTCAGCCAGCAAGCT -GGAAGTCTTCTCTGTGAC 3265  
QY 422 CATGGGCTAAATGGTCTCTCTCTGGAATTTGGGCTTATCAGATAAAACAAGTAGTCA 481  
DB 3264 CATAACTAAATTAATCTCTCTCTGGAATTTGGGCTTATCAGATAAAACAAGTAGTCA 3205  
QY 482 TCCACAGAGATCTCTATAAGCCATGATTTCTGGGATTTCTAGAGTATGCTGATATGAC 541  
DB 3204 TACCACAAATATCTATAAGCCATGATTTCTGGGATTTCTAGAGTATGCTGATATGAC 3145  
QY 542 TAAGCCAGGAGAGCTATTTAAAGATCTCAGCATCTTTCAGCTTTTAACTAGAGAAA 601  
DB 3144 TAAGCCAGGAGAGCTATTTAAAGATCTCAGCATCTTTCAGCTTTTAACTAGAGAAA 3085  
QY 602 ACCCGAAGCATGCTGGATTATAAGGGAATTTGAATGGGTCCACCAA 650  
DB 3084 ACCCGAAGCATGCTGGATTATAAGGGAATTTGAATGGGTCCACCAA 3036

## RESULT 5

US-10-311-455-1625

; Sequence 1625, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ

; TITLE OF INVENTION: cytosine methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1625

; LENGTH: 5926

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1625

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 5704
; OTHER INFORMATION: n is a or g or c or t
;
US-10--311-455-973

Query Match          5.9%; Score 38.2; DB 15; Length 6557;
Best Local Similarity 48.8%; Pred. No. 0.39; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 108;

QY 427 GCTAATTGGTCTCCCTCTCTGGATGTGGCTTATCAGATAAAAAACAAGTGAATCATGCCA 486
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GTTATTGTGATGTGTTTTGGTTTTATACGAAAAATAATAATTAGTTAGGTATGATG 398

QY 487 CAGGATGCTCTAAGCCCAATTGAATCTGGGATCTCTATGATGATGCTGATATGACTTAAGC 546
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GTGGATATTGTAATTTTATTTAGTTATTTTGGAAAGTTTAGGTAGAGAGATCGTTTGAGTTGG 458

QY 547 CAGGACGAGACTTATTTAAAGATCTCAGCATCTTTACGTTGTTTAAACCTAGAGAAACCCG 606
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 GAGTGGAGGTATAGGGAGTTGAGATTATTAATTGTTATTATAGTTTGGTAAATAGAC 518

QY 607 AAGCATGACTGGATTATTAAGGGAATTGAA 637

Db 519 GAGATTTTGTTTTTTAAAAAATAAAAAATAA 549

```

RESULT 7  
US-10-831-704-94  
; Sequence 94, Application US/10831704  
; Publication No. US20050100931A1  
; GENERAL INFORMATION:  
; APPLICANT: Shvjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND  
; TITLE OF INVENTION: ASSESSMENT  
; TITLE OF INVENTION: OF PROSTATE  
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/10/831,704

```

CURRENT FILING DATE: 2004-04-23
PRIOR APPLICATION NUMBER: US/10/155,653
PRIOR FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US/09/220,132
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: US 60/079,303
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: US 60/068,821
PRIOR FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 94
LENGTH: 1497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1497)
OTHER INFORMATION: n = A,T,C or G
US-10-831-704-94

```

[illegible]

Query Match	56.2%;	Score 365.4;	DB 15;	Length 5926;
Best Local Similarity	73.6%;	Pred. No. 2.1e-112;	Indels 1;	Gaps 1;
Matches 479;	Conservative	0;	Mismatches 171;	
QY	1	GGCTGTGTCGAGATGTTCCAGACACAGCCCATGTGAGAGCTCCCTGGCTCGGGGCCA	60	
DB	2241	GGTTTGTGTCGAGATGTTTTAGTATAGTTTAITGTGAGAGTTTTTGGTTTCGGGTTTA	2300	
QY	61	GTATCTGGAATCAGGCTCCAGCCAAATGCAATCTCTCTACGGGACTGGGAATCTCCA	120	
DB	2301	GTATTTGGAATCTAGGTTTTAGTTAAATGTAATTTTTTTTACGGGATTTGGAAATTTTA	2360	
QY	121	AACTGTCCTCCTCAGAGTGGGAATTTCCACTCACTTCTCTCAGCCACACTGACCTCCC	180	
DB	2361	AAAGTTGTTTTTTAGAGTGGGAATTTTTATTTATTTTTTTTTAGTTAGTATGATTTTT	2420	
QY	181	AGCGGGGAGGGGATCTTTTCTTGACAGACAGAGTGGGAGGACAGAGCTGTCACTTT	240	
DB	2421	AGCGGGGAGGGGATTTTTTTTTTGATAGAGTGAAGTGGGAGTAGATGTTATTTATTTT	2480	
QY	241	CGAAGAAGCTTTCTTTCTGATTCATACCCCTTCACTTCCCTGTGTTTACTGTCTGATAT	300	
DB	2481	TTAGAAGATTTTTTTTTTTTGATTTATTTTTTTTTTTTTTGTGTTTATTTGTTTGATAT	2540	
QY	301	ATCCAAAGGCCAAGTCACTTTCCAGAGATGACAACCTCTTCCCTGAAGTAGAGACATGCTT	360	
DB	2541	ATGTAAGGTTAGTTATTTTTTAGAGATGATAATTTTTTTTTTGAAGTAGAGATATGTTT	2600	
QY	361	CCAACACTCAGAAGCCATGTGAACTACTCAGCCAGCAAGCT-GGAAGTTTTTCTGTGTG	419	
DB	2601	TTAATATTTAGAAGTTTTATGTGAATATTTAGTTAGTAAGTTGGGAAGTTTTTTTTTGTG	2660	
QY	420	ACCATGGCTAAATGGTCTCTCTCTCGAATGTGGCTTATCAGATAAAAAACAAGTCACT	479	
DB	2661	ATTATGGGTTAAATGGTTTTTTTTTTTTTTTGGATTTGATTTGTTTATTTAGATAAAATAGTGCT	2720	
QY	480	CATGCCACAGGATGTCATTAAGCCCATTTGATCTGGGATTCATGATGATGCTGATATG	539	
DB	2721	TATGTTATAGATGTTTTATAAGTTTATTTGATTTTGGGATTTTATAGAGTGAATGATG	2780	
QY	540	ACTAAGCCGAGGAGACTTATTTAAAGATCTCAGCATCTTTTCAAGCTTGTTTAACTAGAGA	599	
DB	2781	ATTAAGTTTAGGAGAGATTTATTTANAGTTTTAGTATTTTTTGTGTTTATTTAGAGA	2840	
QY	600	AAACCCGAAGCATCACTGGATTTATAAGGGAAATTTGAATCGGTCACACAA	650	
DB	2841	AAATCTCAGATGATGATTTGATTTAAGGGAATTTGAATCGGTTTATTTAA	2891	

```

RESULT 6
US-10-311-455-973
; Sequence 973, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 973
; LENGTH: 6557
; TYPE: DNA

```

```
QY 145 TTCCACTCACTTCTCTCAGCCAGCACTGACCTCCAGCGGGGAGGCATCTTTCTTG 204
    || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 399 TTWANTTTTMMKTYCMRGSYTKRMRGSCCYKGGGSMWYYWARSWARGRTGSMWW 458
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 ACAGAGCAGAACTGGAGGACAGACAGCTCTCACTTCCAGAAGACTTTCTTTCTGATTC 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 459 AAATGKGYMAWYTYTGKGAARRKTKGGKMYTTTTTKGYTMAWTTTTTMAWTMCCMAWY 518
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 ATACCT 271
    | : : : : :
Db 519 TTYIYK 525
    | : : : : :
RESULT 8
US-10-027-632-254387
; Sequence 254387, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMERIZATION NUMBER: US/10027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254387
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-254387
```

```
Query Match 5.6%; Score 36.2; DB 13; Length 1266;
Best Local Similarity 53.1%; Pred. No. 0.73;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTCGGAACCTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 781 CCATGGGCGGGAGGTGGCAACATAGTGTGATGAGGGGGAGTGGCCCACTGCA 840
QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGAGGCATCTTTCTTGACAGACAGAGTG 218
Db 841 TATCTGCTGGCGCTGAAATCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900
QY 219 GGAGGACAGACGTGTCTCACTTTCCA 243
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925

RESULT 9
US-10-027-632-254389
; Sequence 254389, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMERIZATION NUMBER: US/10027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254389
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-254387
```

```
Query Match 5.6%; Score 36.2; DB 13; Length 1266;
Best Local Similarity 53.1%; Pred. No. 0.73;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTCGGAACCTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 781 CCATGGGCGGGAGGTGGCAACATAGTGTGATGAGGGGGAGTGGCCCACTGCA 840
QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGAGGCATCTTTCTTGACAGACAGAGTG 218
Db 841 TATCTGCTGGCGCTGAAATCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900
QY 219 GGAGGACAGACGTGTCTCACTTTCCA 243
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925
```

```
RESULT 9
US-10-027-632-254388
; Sequence 254388, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
```

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254388
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-254388
```

```
Query Match 5.6%; Score 36.2; DB 13; Length 1266;
Best Local Similarity 53.1%; Pred. No. 0.73;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 99 CTACGGGATCTCGGAACCTTCCAAAGCTGCTCTCAGAGTGGGAATTTCCACTCACTTCT 158
Db 781 CCATGGGCGGGAGGTGGCAACATAGTGTGATGAGGGGGAGTGGCCCACTGCA 840
QY 159 CTCAGCCAGCACTGACCTCCAGCGGGGAGGCATCTTTCTTGACAGACAGAGTG 218
Db 841 TATCTGCTGGCGCTGAAATCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900
QY 219 GGAGGACAGACGTGTCTCACTTTCCA 243
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925
```

```
RESULT 10
US-10-027-632-254389
; Sequence 254389, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMERIZATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254389
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-254389
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254387
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-254387

Query Match          5.6%; Score 36.2; DB 17; Length 1266;
Best Local Similarity 53.1%; Pred. No. 0.73; Indels 0; Gaps 0;
Matches 77; Conservative

Qy 99 CTACGGGATCTGGGAACCTTCCAAAGCTGCTCTCCAGAGTGGGAATTTCCACTCACTTCT 158
Db 781 CCATGGGCGCGGAGGTGGCAACATAGTGCATGAGAGGGGAGCTGCCCCACCTGCA 840
Qy 159 CTCACGCCAGCACTGACCTCCAGCGGGGAGGGGATCTTTCTTCACAGAGCAGAAGTG 218
Db 841 TATCTGCTGGCGCTGAAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAAGATGAGAG 900
Qy 219 GGAGGCAGACAGCTGTCACTTTCCA 243
Db 901 GACTTCACAAAGCTGCCACTGCCA 925

RESULT 13
US-10-027-632-254388
; Sequence 254388, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 254388
;

Query Match          5.6%; Score 36.2; DB 13; Length 1266;
Best Local Similarity 53.1%; Pred. No. 0.73; Indels 0; Gaps 0;
Matches 77; Conservative

Qy 99 CTACGGGATCTGGGAACCTTCCAAAGCTGCTCTCCAGAGTGGGAATTTCCACTCACTTCT 158
Db 781 CCATGGGCGCGGAGGTGGCAACATAGTGCATGAGAGGGGAGCTGCCCCACCTGCA 840
Qy 159 CTCACGCCAGCACTGACCTCCAGCGGGGAGGGGATCTTTCTTCACAGAGCAGAAGTG 218
Db 841 TATCTGCTGGCGCTGAAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAAGATGAGAG 900
Qy 219 GGAGGCAGACAGCTGTCACTTTCCA 243
Db 901 GACTTCACAAAGCTGCCACTGCCA 925

RESULT 12
US-10-027-632-254387
; Sequence 254387, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:

```

; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-254388

Query Match 5.6%; Score 36.2; DB 17; Length 1266;  
Best Local Similarity 53.1%; Pred. No. 0.73;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 99 CTACGGGATCTGGGAACCTTCCAAAGCTGCCTCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 781 CCATGGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCCACTTCTGCA 840  
  
QY 159 CTCACGCCACGACTGACCTCCAGCGGGGAGGGGCATCTTTTTCACAGAGCAGAAAGTG 218  
Db 841 TATCTGCTGGCGCTGAAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900  
  
QY 219 GGAGGCAGACAGCTGTCTCTTTCCA 243  
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925

## RESULT 14

US-10-027-632-254389  
; Sequence 254389, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: US 60/146,002

; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 254389  
; LENGTH: 1266

; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-254389

Query Match 5.6%; Score 36.2; DB 17; Length 1266;  
Best Local Similarity 53.1%; Pred. No. 0.73;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 99 CTACGGGATCTGGGAACCTTCCAAAGCTGCCTCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 781 CCATGGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCCACTTCTGCA 840  
  
QY 159 CTCACGCCACGACTGACCTCCAGCGGGGAGGGGCATCTTTTTCACAGAGCAGAAAGTG 218  
Db 841 TATCTGCTGGCGCTGAAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900  
  
QY 219 GGAGGCAGACAGCTGTCTCTTTCCA 243  
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925

## RESULT 15

US-10-027-632-254390  
; Sequence 254390, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 254390  
; LENGTH: 1266  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-254390

Query Match 5.6%; Score 36.2; DB 17; Length 1266;  
Best Local Similarity 53.1%; Pred. No. 0.73;  
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
  
QY 99 CTACGGGATCTGGGAACCTTCCAAAGCTGCCTCTCAGAGTGGGAATTTCCACTCACTTCT 158  
Db 781 CCATGGGGCGGGAGGTGGCAACATAGTGCATGAGAGGGGACTGCCCCCACTTCTGCA 840  
  
QY 159 CTCACGCCACGACTGACCTCCAGCGGGGAGGGGCATCTTTTTCACAGAGCAGAAAGTG 218  
Db 841 TATCTGCTGGCGCTGAAATTTCTCTCAGGACAGCTCAGCCCTGCTTTCCAGAGATGAGAG 900  
  
QY 219 GGAGGCAGACAGCTGTCTCTTTCCA 243  
Db 901 GACTTCACAAAGCTGCCCACTGCCA 925

Search completed: August 4, 2005, 15:54:28  
Job time : 1161.47 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:53:02 ; Search time 80.3887 Seconds  
(without alignments)  
12055.236 Million cell updates/sec

Title: US-09-899-276C-8  
Perfect score: 20  
Sequence: 1 ggaagggttgagtcaggatt 20

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	AX343335 Sequence
2	18.4	92.0	227648	10	AC121286 Mus muscu
3	18.4	92.0	234354	10	AC116406 Mus muscu
4	18	90.0	155753	9	AC104259 Homo sapi
5	18	90.0	164359	9	AC009899 Homo sapi
6	18	90.0	177456	2	AP001356 Homo sapi
7	17.4	87.0	2172	10	AF071772 Mus muscu
8	17.4	87.0	102448	10	AL928696 Mouse DNA
9	17.4	87.0	140334	9	AC021802 Homo sapi
10	17.4	87.0	189983	2	AC133524 Mus muscu
11	17.4	87.0	191426	2	AC026086 Homo sapi
12	17.4	87.0	193363	10	AL627104 Mouse DNA
13	17.4	87.0	194546	9	AC087482 Homo sapi
14	17.4	87.0	208945	2	AC125890 Rattus no
15	17.4	87.0	224855	2	AC095610 Rattus no
16	17.4	87.0	234805	2	AC112287 Rattus no
17	17.4	87.0	246721	2	AC110714 Rattus no
18	17.4	87.0	248646	2	AC113650 Rattus no
19	17.4	87.0	268965	2	AC094228 Rattus no

C 20	17.4	87.0	276300	2	AC127855
C 21	17.4	87.0	295086	2	AC095195
C 22	17.4	87.0	302176	2	AC096479
C 23	17	85.0	21	6	AX343337 Sequence
C 24	17	85.0	600	6	AX343328 Sequence
C 25	17	85.0	9174	9	AF519531 Homo sapi
C 26	17	85.0	11793	9	HSY18933 Y18933 Homo sapi
C 27	17	85.0	147416	9	AC005549 Homo sapi
C 28	17	85.0	245273	2	AC137407 Rattus no
C 29	17	85.0	252902	2	AC129164 Rattus no
C 30	17	85.0	268213	2	AC103285 Rattus no
C 31	16.8	84.0	600	14	AF510258 Human rse
C 32	16.8	84.0	1274	8	BT013955 Lycopersi
C 33	16.8	84.0	1575	8	BT005812 Arabidops
C 34	16.8	84.0	1772	8	AY094410 Arabidops
C 35	16.8	84.0	2263	4	RABLINE1E M93325 Rabbit SINE
C 36	16.8	84.0	2310	8	SPWIS1 X62631 S.pombe wls
C 37	16.8	84.0	2511	9	AK025126 Homo sapi
C 38	16.8	84.0	10249	10	AF345640 Mus muscu
C 39	16.8	84.0	16695	5	AP004103 Esoc luci
C 40	16.8	84.0	16743	4	AB096865 Procavia
C 41	16.8	84.0	4865	8	SPSC409 AL109822 S.pombe c
C 42	16.8	84.0	45701	9	AC099412 Homo sapi
C 43	16.8	84.0	48173	9	AL583849 Human DNA
C 44	16.8	84.0	59356	9	AC115109 Homo sapi
C 45	16.8	84.0	76738	9	AL138813 Human DNA

ALIGNMENTS

RESULT 1  
AX343335  
LOCUS AX343335  
DEFINITION Sequence 8 from Patent EP1170372.  
ACCESSION AX343335  
VERSION AX343335.1 GI:18491685  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Roessl,F., Soto,U., Coy,J., Finzer,P., Delius,H., Poustka,A., zur Hausen,H. and Patzelt,A.  
TITLE Regulatory sequences of the human mcp-1 gene  
JOURNAL Patent: EP 1170372-A 8 09-JAN-2002;  
DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES OEFFENTLICHEN RECHTS (DE)  
FEATURES  
Location/Qualifiers  
1..20  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description of the Artificial Sequence: Part of 3'-DHSR"

ORIGIN	Query Match	100.0%;	Score 20;	DB 6;	Length 20;
Best Local Similarity	Matches	20;	Conservative	0;	Mismatches
0;	Indels	0;	Gaps	0;	
QY	1	GGAAGGTTGAGTCAAGGATT	20		
Db	1	GGAAGGTTGAGTCAAGGATT	20		
RESULT 2	AC121286/c	227648 bp	DNA	linear	ROD 06-MAR-2004
LOCUS AC121286	Mus musculus chromosome 12, clone RP23-36H21, complete sequence.				
DEFINITION AC121286					
ACCESSION AC121286					
VERSION AC121286.7					
KEYWORDS HTG.					

mus musculus (house mouse)	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 (bases 1 to 227648)	
Birken,B., Nusbaum,C. and Lander,E.	
Mus musculus chromosome 12, clone RP23-36H21	
Unpublished	
2 (bases 1 to 227648)	
Birken,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazato,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,B., Grahm,U., Johnson,R., Jones,C., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rottet,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
3 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., McLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
4 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
5 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
6 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
7 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
8 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramaamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
Direct Submission	
Submitted (22-JAN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
9 (bases 1 to 227648)	
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,U., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W.,	

repeat\_region /rpt\_family="(TC)n"  
10646 .10898  
repeat\_region /rpt\_family="Lx9"  
11949 .11975  
repeat\_region /rpt\_family="(TC)n"  
12159 .12198  
repeat\_region /rpt\_family="(T)n"  
12202 .12374  
repeat\_region /rpt\_family="L1\_MM"  
12375 .12415  
repeat\_region /rpt\_family="AT\_rich"  
12576 .12729  
repeat\_region /rpt\_family="(GAAA)n"  
12754 .12812  
repeat\_region /rpt\_family="(TAAAA)n"  
12843 .12926  
repeat\_region /rpt\_family="(TTCTC)n"  
complement(12943) .13007  
repeat\_region /rpt\_family="Lx"  
13001 .13290  
repeat\_region /rpt\_family="L1"  
13295 .13321  
repeat\_region /rpt\_family="(TTTG)n"  
complement(13334) .13478  
repeat\_region /rpt\_family="B1\_MM"  
complement(13561) .13582  
repeat\_region /rpt\_family="B1\_MM"  
complement(13583) .13648  
repeat\_region /rpt\_family="MT2B"  
complement(13665) .13725  
repeat\_region /rpt\_family="ORR1A-int"  
13779 .13800  
repeat\_region /rpt\_family="(TTTTTG)n"  
complement(13816) .14126  
repeat\_region /rpt\_family="Lx2"  
complement(14121) .14515  
repeat\_region /rpt\_family="Lx4"  
14519 .14624  
repeat\_region /rpt\_family="Lx8"  
15081 .16704  
repeat\_region /rpt\_family="L1\_MM"  
18107 .18757  
repeat\_region /rpt\_family="L1MAS"  
19079 .19134  
repeat\_region /rpt\_family="(GA)n"  
19138 .20151  
repeat\_region /rpt\_family="Lx8"  
20999 .21032  
repeat\_region /rpt\_family="(T)n"

Query Match 92.0%; Score 18.4; DB 10; Length 227648;  
Best Local Similarity 95.0%; Pred.No. 44;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20  
|||||  
Db 202278 GGAAGGTTGACTCAAGGATT 202259

RESULT 3  
AC116406/c AC116406 234354 bp DNA linear ROD 27-SEP-2003  
DEFINITION Mus musculus chromosome 3, clone RP23-427K12, complete sequence.  
ACCESSION AC116406  
VERSION AC116406.12 GI:34581725  
KEYWORDS HTG.

SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 234354)  
Birken,B., Nusbaum,C., Lander,E.  
TITLE Mus musculus chromosome 3, clone RP23-427K12

JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 234354)  
Birken,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,  
Landers,T., Lechoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McGwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (27-MAR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 234354)  
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,  
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,  
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,  
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,  
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,  
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (02-AUG-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 234354)  
Birken,B., Nusbaum,C. and Lander,E.

JOURNAL  
REFERENCE  
AUTHORS

Submitted (12-SEP-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
5 (bases 1 to 234354)  
Birken,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,  
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,  
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,  
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,  
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,

Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
 Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,  
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (27-SEP-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 12, 2003 this sequence version replaced gi:33413361.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
 Center project name: L23945  
 Center clone name: 427\_K\_12  
 -----

## FEATURES

## Location/Qualifiers

1. .234354  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="3"  
 /map="3"  
 /clone="RP23-427K12"  
 /clone\_lib="RP23-427K12 Female Mouse BAC"

1. .32659  
 /notes="wgs end extension  
 clone end:SPC"

3. .32  
 /rpt\_family="(GA)n"  
 complement(1629..1854)

/rpt\_family="B4A"  
 complement(2115..2441)

/rpt\_family="ORR1A3"  
 2895..3003

/rpt\_family="RMER15"  
 3855..3901

/rpt\_family="(CA)n"  
 4188..4224

/rpt\_family="MER5A"  
 5391..5414

/rpt\_family="(TA)n"  
 complement(5560..5690)

/rpt\_family="B1\_MM"  
 6972..7020

/rpt\_family="(TC)n"  
 7911..7988

/rpt\_family="ID3"  
 complement(8362..8647)

/rpt\_family="B4A"  
 8741..8851

/rpt\_family="(TTTC)n"  
 8936..8966

/rpt\_family="(TCTG)n"  
 8987..9009

/rpt\_family="(TTCC)n"  
 9550..9697

/rpt\_family="L1MB5"  
 complement(9754..10062)

/rpt\_family="ORR1A2"  
 10215..10257

/rpt\_family="GA-rich"  
 complement(10436..10797)

/rpt\_family="ORR1C"  
 10798..11102

/rpt\_family="ORR1A3"  
 complement(11103..11129)

/rpt\_family="ORR1C"  
 complement(11155..11513)

repeat\_region  
 /rpt\_family="Lx9"  
 12003..12023  
 /rpt\_family="AT-rich"  
 12178..12220  
 /rpt\_family="T-rich"  
 12731..12760  
 /rpt\_family="(TTTTA)n"  
 complement(12848..13133)  
 /rpt\_family="B4A"  
 13261..13403  
 /rpt\_family="MER5A"  
 complement(13906..14971)  
 /rpt\_family="Lx8"  
 16202..16531  
 /rpt\_family="MER34B-int"  
 17055..17241  
 /rpt\_family="B2\_Mm1"  
 17513..18294  
 /rpt\_family="MER34B-int"  
 18411..18709  
 /rpt\_family="MER31B"  
 18753..18901  
 /rpt\_family="RSINE1"  
 18920..19075  
 /rpt\_family="(TAGA)n"  
 complement(19834..19912)  
 /rpt\_family="PBID9"  
 complement(19927..20155)  
 /rpt\_family="B3"  
 20579..20631  
 /rpt\_family="(TTTA)n"  
 20633..20711  
 /rpt\_family="(TG)n"  
 complement(20716..21003)  
 /rpt\_family="B4"  
 21249..21311  
 /rpt\_family="A-rich"  
 21880..21957  
 /rpt\_family="(TA)n"  
 complement(21975..22067)  
 /rpt\_family="RSINE1"  
 complement(22294..22371)

Query Match 92.0%; Score 18.4; DB 10; Length 234354;  
 Best Local Similarity 95.0%; Pred. No. 44;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAGGTTGAGTCAGGATT 20

Db 107064 GGAGGTTGAGTCAGGATT 107045

## RESULT 4

AC104259/c

LOCUS

DEFINITION

AC104259

VERSION

HTG.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC104259 155753 bp DNA linear PRI 22-JAN-2003  
 Homo sapiens chromosome 15, clone CTD-252514, complete sequence.  
 AC104259  
 AC104259.14 GI:27819574  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 155753)  
 Birren, B., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 15, clone CTD-252514  
 Unpublished  
 2 (bases 1 to 155753)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# TITLE JOURNAL

Submitted (06-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 155753)

## REFERENCE AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# TITLE JOURNAL

Submitted (21-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 155753)

## REFERENCE AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# TITLE JOURNAL

Submitted (22-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jan 22, 2003 this sequence version replaced gi:25141082.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Repeat code: WIBR

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L22081  
Center clone name: 2525\_1\_4  
-----

Only the first 155.8 kilobases of this clone are being submitted.  
The remainder overlaps accession number AC013489 [WICGR project L1674].

## FEATURES

source	Location/Qualifiers
1. .155753	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="15"
	/map="15"
	/clone="CTD-252514"
	/clone_lib="CITD2 Human BAC"
	/complement(20..148)
repeat_region	/rpt_family="MIR3"
repeat_region	/complement(1166..1215)
repeat_region	/rpt_family="MER5A"
repeat_region	1219..1872
repeat_region	/rpt_family="LIPB3"
repeat_region	1906..2031
repeat_region	/rpt_family="LIPB3"
repeat_region	2032..2056
repeat_region	/rpt_family="AT_rich"
repeat_region	2057..2381
repeat_region	/rpt_family="AluSq"
repeat_region	/complement(2383..2504)
repeat_region	/rpt_family="MER5A"
repeat_region	/complement(2541..2615)
repeat_region	/rpt_family="L2"
repeat_region	/complement(2628..2770)
repeat_region	/rpt_family="Charlie4"
repeat_region	2793..2815
repeat_region	/rpt_family="(CAAAAA)n"
repeat_region	/complement(2822..3166)
repeat_region	/rpt_family="L2"
repeat_region	3306..3406
repeat_region	/rpt_family="MIR3"
repeat_region	3649..3673
repeat_region	/rpt_family="GC_rich"
repeat_region	3692..3812
repeat_region	/rpt_family="C-rich"
repeat_region	3890..3924
repeat_region	/rpt_family="GC_rich"
repeat_region	/complement(4145..4277)
repeat_region	/rpt_family="L2"
repeat_region	/complement(4292..4395)
repeat_region	/rpt_family="MIR"
repeat_region	4472..5119
repeat_region	/rpt_family="LIME3A"
repeat_region	5160..5397
repeat_region	/rpt_family="LIM4"
repeat_region	5514..5690
repeat_region	/rpt_family="LIM4"
repeat_region	5699..5986
repeat_region	/rpt_family="AluJb"
repeat_region	5988..6009
repeat_region	/rpt_family="(CA)n"
repeat_region	6229..6747
repeat_region	/rpt_family="LIME1"
repeat_region	/complement(6912..7004)
repeat_region	/rpt_family="MIR"
repeat_region	9197..9328
repeat_region	/rpt_family="MIR3"
repeat_region	10922..10965
repeat_region	/rpt_family="MIR"
repeat_region	11022..11114
repeat_region	/rpt_family="L2"

Tue Aug 9 17:01:29 2005

```

repeat_region      11651..12118
                    /rpt family="L3"
repeat_region      complement(12943..13015)
                    /rpt family="MIR3"
repeat_region      13736..13768 (TG) n
                    /rpt family="MIR"
repeat_region      14496..14560
                    /rpt family="MIR"
repeat_region      14561..14610
                    /rpt family="MLT1P"
repeat_region      14875..15063
                    /rpt family="MIR"
repeat_region      complement(16061..16132)
                    /rpt family="LIM4"
repeat_region      16133..16424
                    /rpt family="AluSx"
repeat_region      complement(16425..16781)
                    /rpt family="LIM4"
repeat_region      16919..16933
                    /rpt family="MER5C"
repeat_region      complement(16934..17221)
                    /rpt family="AluJb"
repeat_region      17222..17393
                    /rpt family="MER5C"
repeat_region      17453..17606
                    /rpt family="MER103"
repeat_region      complement(17669..17767)
                    /rpt family="L3"
repeat_region      17882..17997
                    /rpt family="L2"
repeat_region      18194..18519
                    /rpt family="L2"
repeat_region      19131..19458
                    /rpt family="L2"
repeat_region      complement(19567..19662)

Query Match      90.0% Score 18; DB 9; Length 155753;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GAAGTTTGAGTCAAGGAT 19
        |||||
Db      31307 GAAGTTTGAGTCAAGGAT 31290

RESULT 5
AC009899      164359 bp      DNA      linear      PRI 24-AUG-2002
LOCUS      AC009899      Homo sapiens chromosome 18, clone RP11-19P9, complete sequence.
DEFINITION      AC009899      HTG.
ACCESSION      AC009899
VERSION      AC009899.12      GI:22475026
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164359)
Birren,B., Nusbaum,C. and Lander,E.
Unpublished
2 (bases 1 to 164359)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cook,P., Dearellano,K., Depayre,B., Fitzhugh,W., Forrest,C.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Grant,G.,
Hague,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,

```

TITLE  
JOURNALREFERENCE  
AUTHORS

Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,  
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.

Submitted (05-SEP-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 164359)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
Murphy,T., Naylor,J., O'Neill,D., Oliver,J., Peterson,K.,  
O'Connor,T., O'Donnell,P., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Phunkhang,P., Pierre,N., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNALREFERENCE  
AUTHORS

Submitted (22-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 164359)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Norbu,C., Norman,C.H.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Peterson,K.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

## COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 24, 2002 this sequence version replaced gi:22417381.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L2244  
Center clone name: 19\_F\_9  
-----

FEATURES  
source

We are submitting the first 164.4 kilobases of the project at this  
time. The remainder overlaps AC091495 [WICGR project LI2630].

Location/Qualifiers  
1..164359  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"

```

/clone="RP11-19F9"
repeat_region 291.468 /rpt_family="MIR"
repeat_region 1063.1288 /rpt_family="MIR"
repeat_region complement(1516.1912) /rpt_family="MLT1H"
repeat_region 2086.2108 /rpt_family="AT_rich"
repeat_region complement(2226.2308) /rpt_family="L2"
repeat_region 2766.2869 /rpt_family="L1MA8"
repeat_region 2867.2965 /rpt_family="L1MA9"
repeat_region 3028.3191 /rpt_family="L1MA9"
repeat_region 3197.3247 /rpt_family="L1MA9"
repeat_region complement(3506.3629) /rpt_family="MIR"
repeat_region 4113.4278 /rpt_family="L3"
repeat_region complement(4392.4456) /rpt_family="L3"
repeat_region 5497.5564 /rpt_family="L2"
repeat_region complement(5567.5706) /rpt_family="L2"
repeat_region 5915.5956 /rpt_family="L2"
repeat_region complement(6312.6587) /rpt_family="L2"
repeat_region 6588.6612 /rpt_family="L2"
repeat_region complement(6613.7883) /rpt_family="L2"
repeat_region complement(9865.10148) /rpt_family="L2"
repeat_region complement(10152.10626) /rpt_family="L2"
repeat_region 1123.11175 /rpt_family="L2"
repeat_region 11309.11376 /rpt_family="L2"
repeat_region 11565.11616 /rpt_family="L2"
repeat_region 12503.12584 /rpt_family="L2"
repeat_region 12587.12643 /rpt_family="L2"
repeat_region complement(12644.12850) /rpt_family="L2"
repeat_region 12905.13274 /rpt_family="L2"
repeat_region complement(13410.13461) /rpt_family="L2"
repeat_region 14051.14225 /rpt_family="L2"
repeat_region 14438.14647 /rpt_family="L2"
repeat_region complement(14748.15152) /rpt_family="L2"
repeat_region 17893.17987 /rpt_family="L2"
repeat_region complement(18333.18426) /rpt_family="L2"
repeat_region 20976.20981 /rpt_family="L2"
repeat_region note="<30 qual SNGL region" /rpt_family="L2"
repeat_region complement(22730.22896) /rpt_family="L2"
repeat_region complement(22930.23337) /rpt_family="L2"

/rpt_family="MLT1H"
23443.23487 /rpt_family="MLT1H"
complement(23513.23779) /rpt_family="L2"
27703.27736 /rpt_family="L2"
27773.27812 /rpt_family="L2"
28640.28671 /rpt_family="L2"
29132.30007 /rpt_family="L2"
complement(30295.32620) /rpt_family="L2"
complement(32622.33185) /rpt_family="L2"
33158.34257 /rpt_family="L2"
34258.34523 /rpt_family="L2"
34524.36097 /rpt_family="L2"
36194.36328 /rpt_family="L2"
36345.36422 /rpt_family="L2"
36428.37415 /rpt_family="L2"

Query Match 90.0%; Score 18; DB 9; Length 164359;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAGGAT 19
DB 83034 GAAGTTGAGTCAGGAT 83051

RESULT 6
AP001356/c
LOCUS Homo sapiens chromosome 18 clone RP11-776M20 map 18q12, WORKING
DEFINITION DRAFT SEQUENCE, 8 unordered pieces.
ACCESSION AP001356
VERSION AP001356.3 GI:9229949
KEYWORDS HTGS, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
REFERENCE 1 (bases 1 to 177456)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 177,456 Genomic DNA of 18q12
PUBLISHED Only in Database (2000)
REFERENCE 2 (bases 1 to 177456)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (06-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 15, 2000 this sequence version replaced gi:8117270.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-776M20

```

----- Summary Statistics -----  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 174936 bases at least Q40  
Consensus quality: 176005 bases at least Q30  
Consensus quality: 176462 bases at least Q20  
Insert size: 176756; sum-of-contigs  
Quality coverage: 11.10x in Q20 bases; sum-of-contigs

-----  
NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 44797 contig of 44797 bp in length  
44898 77452 contig of 32555 bp in length  
77553 105748 contig of 28196 bp in length  
105849 132030 contig of 26182 bp in length  
132131 154722 contig of 22592 bp in length  
154823 165228 contig of 10406 bp in length  
165329 172368 contig of 7040 bp in length  
172469 177456 contig of 4988 bp in length.  
\* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 44797: contig of 44797 bp in length  
44898 77452: contig of 32555 bp in length  
77553 105748: contig of 28196 bp in length  
105849 132030: contig of 26182 bp in length  
132131 154722: contig of 22592 bp in length  
154823 165228: contig of 10406 bp in length  
165329 172368: contig of 7040 bp in length  
172469 177456: contig of 4988 bp in length.

FEATURES

source  
1..177456  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18q12"  
/clone="RP11-776M20"  
1..44797  
/note="assembly\_fragment"  
44898..77452  
/note="assembly\_fragment"  
77553..105748  
/note="assembly\_fragment clone\_end:SP6 vector\_side:left"  
105849..132030  
/note="assembly\_fragment"  
132131..154722  
/note="assembly\_fragment"  
154823..165228  
/note="assembly\_fragment clone\_end:T7 vector\_side:right"  
165329..172368  
/note="assembly\_fragment"  
172469..177456  
/note="assembly\_fragment"

ORIGIN

Query Match 90.0%; Score 18; DB 2; Length 177456;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 GAAGGTTGAGTCAAGGAT 19  
|||||  
Db 142571 GAAGGTTGAGTCAAGGAT 142554

RESULT 7

AF071772 2172 bp DNA linear ROD 22-SEP-1998  
LOCUS Mus musculus beta2/neuroD gene, promoter and exon 1.  
DEFINITION AF071772  
ACCESSION AF071772  
VERSION AF071772.1 GI:3641530  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2172)  
AUTHORS Xu, W. and Murphy, L.J.  
TITLE Isolation and characterization of the mouse beta 2/neuroD gene promoter  
JOURNAL Biochem. Biophys. Res. Commun. 247 (3), 814-818 (1998)  
MEDLINE 98321210  
PUBMED 9647776  
REFERENCE 2 (bases 1 to 2172)  
AUTHORS Xu, W. and Murphy, L.J.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUN-1998) Physiology, University of Manitoba, 770 Bannatyne Ave., Winnipeg, MB R3E 0W3, Canada  
FEATURES  
Location/Qualifiers  
1..2172  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6N"  
/db\_xref="taxon:10090"  
/chromosome="2"  
1..2090  
2060..2066  
2091..2172  
/product="beta2/neuroD"  
/note="helix-loop-helix protein"  
2091..2172  
/number=1  
2091..2172  
5'UTR  
ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 2172;  
Best Local Similarity 94.7%; Pred. No. 3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGAT 19  
|||||  
Db 1307 GGAAGGTTGAGTCAAGGCT 1325

RESULT 8

AL928696/c 102448 bp DNA linear ROD 21-NOV-2002  
LOCUS Mouse DNA sequence from clone RP23-387G11 on chromosome 2, complete  
DEFINITION AL928696  
ACCESSION AL928696  
VERSION AL928696.6 GI:25168716  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 102448)  
AUTHORS Sycamore, N.

**TITLE**  
JOURNAL

**COMMENT**

Direct Submission  
Submitted (21-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 22, 2002 this sequence version replaced gi:25136705.  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a phred quality of at least 30.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquery@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPSP; Information  
on the WORMPSP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) RP23-387G11 is  
from the RPCI-23 Mouse PAC library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACe3.6

**FEATURES**  
Source

1..102448  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosomes="2"  
/clone="RP23-387G11"  
/clone\_lib="RPCI-23"

**ORIGIN**

Query Match 87.0%; Score 17.4; DB 10; Length 102448;  
Best Local Similarity 94.7%; Pred. NO. 1.7e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAGGTTGAGTCACGAT 19  
|||||  
DB 6386 GGAAGGTTGAGTCAGGCT 6368

**RESULT 9**  
AC021802/c  
LOCUS AC021802 140334 bp DNA linear PRI 01-DEC-2001  
DEFINITION Homo sapiens, clone RP11-28C3, complete sequence.  
ACCESSION AC021802  
VERSION AC021802.6 GI:16974240  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 140334)  
REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLES Homo sapiens, clone RP11-28C3  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 140334)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,  
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkert, G., Castle, A.,  
Chospel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,  
DeArelano, K., Dewar, K., Domino, M., Doyle, M., Fenster, J.,  
Ferreira, P., Fitch, H., W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lander, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
Zimmer, A. and Zody, M.

## TITLE

Direct Submission  
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 140334)

## REFERENCE

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,  
Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,  
Chospel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
Ferreira, P., Fitch, H., W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K.,  
Lamazar, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Olivari, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission  
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 18, 2001 this sequence version replaced gi:15706157.

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIER  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L4744  
Center clone name: 28\_C\_3

## FEATURES

Source  
1..140334  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-28C3"  
/clone\_lib="RPCI-11 Human Male BAC"  
complement (502..677)  
repeat\_region /rpt\_family="MIR"  
820..885  
repeat\_region /rpt\_family="MIR"  
1188..1247  
repeat\_region /rpt\_family="MIR"  
2104..2105  
unsure

repeat_region	/note="single clone coverage"	repeat_region	13274..13406
repeat_region	2395..2596	repeat_region	/rpt_family="L2"
repeat_region	/rpt_family="MIR"	repeat_region	14009..14052
repeat_region	2801..2852	repeat_region	/rpt_family="C-rich"
repeat_region	/rpt_family="L1MB3"	repeat_region	14071..14118
repeat_region	complement(2853..3158)	repeat_region	/rpt_family="(TG)n"
repeat_region	/rpt_family="AluSp"	repeat_region	14804..14864
repeat_region	3159..3508	repeat_region	/rpt_family="MIR"
repeat_region	/rpt_family="L1MB3"	repeat_region	15850..16050
unsure	3416..3422	repeat_region	/rpt_family="MIR3"
unsure	/note="qual <30"	repeat_region	16152..16298
repeat_region	3486..3488	repeat_region	/rpt_family="MIR"
unsure	/note="qual <30"	repeat_region	complement(16584..16875)
repeat_region	complement(3509..3810)	repeat_region	/rpt_family="AluJb"
unsure	/rpt_family="AluSq"	repeat_region	17230..17525
repeat_region	3547..3552	repeat_region	/rpt_family="L1MB3"
repeat_region	3811..4197	repeat_region	17552..17683
repeat_region	/rpt_family="L1MB3"	repeat_region	/rpt_family="L1MB3"
repeat_region	4207..4503	repeat_region	17684..17994
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="AluSx"
repeat_region	5963..5998	repeat_region	17995..18764
repeat_region	/rpt_family="AT-rich"	repeat_region	/rpt_family="L1MB3"
repeat_region	6040..6342	repeat_region	18853..18898
unsure	/rpt_family="AluSq"	repeat_region	/rpt_family="L1MB3"
repeat_region	complement(6288..6322)	repeat_region	19089..19215
repeat_region	/note="single clone coverage"	repeat_region	/rpt_family="AluJb"
repeat_region	6351..6533	repeat_region	19216..19541
repeat_region	/rpt_family="L1MB3"	repeat_region	/rpt_family="AluSc"
repeat_region	complement(6534..6819)	repeat_region	19542..19705
repeat_region	/rpt_family="AluY"	repeat_region	/rpt_family="AluJb"
repeat_region	6820..6953	repeat_region	complement(20485..20550)
repeat_region	/rpt_family="L1MB3"	repeat_region	/rpt_family="MIR"
repeat_region	6954..7259	repeat_region	20551..20858
repeat_region	/rpt_family="AluSx"	repeat_region	/rpt_family="AluSp"
repeat_region	7260..7489	repeat_region	complement(20859..20934)
repeat_region	/rpt_family="L1MB3"	repeat_region	/rpt_family="MIR"
repeat_region	7567..7882	repeat_region	21107..21398
repeat_region	/rpt_family="AluSc"	repeat_region	/rpt_family="AluJb"
repeat_region	complement(8219..8448)	repeat_region	21400..21637
repeat_region	/rpt_family="MIR"		/rpt_family="MIR"
repeat_region	complement(8500..8603)		
repeat_region	/rpt_family="L1MB4A"		
repeat_region	complement(8604..8905)		
repeat_region	/rpt_family="AluJo"		
repeat_region	complement(8906..9064)		
repeat_region	/rpt_family="L1MB4A"		
repeat_region	9386..9953		
repeat_region	/rpt_family="L2"		
repeat_region	complement(9954..10055)		
repeat_region	/rpt_family="MER33"		
repeat_region	10056..10102		
repeat_region	/rpt_family="(TA)n"		
repeat_region	complement(10103..10410)		
repeat_region	/rpt_family="AluSg"		
repeat_region	complement(10411..10618)		
repeat_region	/rpt_family="MER33"		
repeat_region	10619..10737		
repeat_region	/rpt_family="L2"		
repeat_region	10997..11136		
repeat_region	/rpt_family="L2"		
repeat_region	11160..11439		
repeat_region	/rpt_family="AluSx"		
repeat_region	11443..11783		
repeat_region	/rpt_family="AluY"		
repeat_region	complement(11784..11894)		
repeat_region	/rpt_family="MIR"		
repeat_region	complement(12227..12469)		
repeat_region	/rpt_family="AluSg"		
repeat_region	complement(12470..12604)		
repeat_region	/rpt_family="AluSq/x"		
repeat_region	12710..12770		
repeat_region	/rpt_family="MIR3"		
Query Match 87.0%; Score 17.4; DB 9; Length 140334;			
Best Local Similarity 94.7%; Pred. No. 1.6e+02;			
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	2 GAAGCTTGAGTCAGGATT 20		
Db	24798 GAAGCTTGAGTCAGGATT 24780		
RESULT 10			
AC133524/c			
LOCUS			
DEFINITION			
AC133524 189983 bp DNA linear HTG 14-SEP-2002			
Mus musculus chromosome UNK clone RP23-43L4, WORKING DRAFT			
SEQUENCE, 5 unordered pieces.			
ACCESSION			
AC133524.1 GI:22830537			
VERSION			
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTROP.			
KEYWORDS			
SOURCE			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 189983)			
McPherson,J.D. and Waterston,R.H.			
TITLE			
The sequence of Mus musculus clone			
JOURNAL			
Unpublished			
REFERENCE			
2 (bases 1 to 189983)			
McPherson,J.D. and Waterston,R.H.			
AUTHORS			
Direct Submission			
TITLE			
Submitted (13-SEP-2002) Genome Sequencing Center, 4444 Forest Park			
JOURNAL			
Parkway, St. Louis, MO 63108, USA			
REFERENCE			
3 (bases 1 to 189983)			
McPherson,J.D. and Waterston,R.H.			
AUTHORS			

TITLE Direct Submission  
JOURNAL Submitted (14-SEP-2002) Genome Sequencing Center, 4444 Forest Park  
COMMENT Parkway, St. Louis, MO 63108, USA

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
----- Project Information -----  
Center project name: M\_BA0043L04

----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-terminator Big Dye; 100%  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 186788 bases at least Q40  
Consensus quality: 187326 bases at least Q30  
Consensus quality: 187762 bases at least Q20  
Insert size: BAC96; agarose-fp  
Insert size: 190782; sum-of-contigs  
Quality coverage: 12.95 in Q20 bases; agarose-fp  
Quality coverage: 9.71 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1969: contig of 1969 bp in length  
\* 1970 2069: gap of unknown length  
\* 2070 5013: contig of 2944 bp in length  
\* 5014 5113: gap of unknown length  
\* 5114 11567: contig of 6454 bp in length  
\* 11568 11668: gap of unknown length  
\* 11668 60663: contig of 48996 bp in length  
\* 60664 60763: gap of unknown length  
\* 60764 189983: contig of 129220 bp in length.

## FEATURES

## source

1. .189983  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP23-43L4"  
1. .1969  
/note="assembly\_name:Contig10"  
2070. .5013  
/note="assembly\_name:Contig12"  
5114. .11567  
/note="assembly\_name:Contig13"  
11668. .60663  
/note="assembly\_name:Contig14"  
60764. .189983  
/note="assembly\_name:Contig15"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 189983;  
Best Local Similarity 94.7%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGAAGGTTGAGTCAAGCAT 19  
|||||  
Db 32815 GGAAGGTTGAGTCAAGGCT 32797

## RESULT 11

AC026086/c

LOCUS AC026086 191426 bp DNA linear HTG 01-SEP-2000  
DEFINITION Homo sapiens chromosome 15 clone RP11-540N15, WORKING DRAFT  
SEQUENCE, 26 unordered pieces.  
ACCESSION AC026086  
VERSION AC026086.4 GI:9958281  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 191426)  
AUTHORS Waterston,R.H.  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 191426)  
TITLE Waterston,R.H.  
JOURNAL Direct Submission  
AUTHORS Submitted (19-MAR-2000) Genome Sequencing Center, Washington  
TITLE University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
JOURNAL MO 63108, USA  
COMMENT On Sep 1, 2000 this sequence version replaced gi:85688910.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H\_NH0540N15  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Chemistry: Dye-terminator Big Dye; 100%  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 180224 bases at least Q40  
Consensus quality: 183561 bases at least Q30  
Consensus quality: 185399 bases at least Q20  
Insert size: 203000; agarose-fp  
Insert size: 188926; sum-of-contigs  
Quality coverage: 4.00 in Q20 bases; agarose-fp  
Quality coverage: 4.34 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 26 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1754: contig of 1754 bp in length  
\* 1755 1854: gap of unknown length  
\* 1855 4193: contig of 2339 bp in length  
\* 4194 4293: gap of unknown length  
\* 4294 6143: contig of 1850 bp in length  
\* 6144 6243: gap of unknown length  
\* 6244 7809: contig of 1566 bp in length  
\* 7810 7909: gap of unknown length  
\* 7910 9758: contig of 1849 bp in length  
\* 9759 9858: gap of unknown length  
\* 9859 15360: contig of 5502 bp in length  
\* 15361 15460: gap of unknown length  
\* 15461 18198: contig of 2738 bp in length  
\* 18199 18298: gap of unknown length  
\* 18299 21882: contig of 3584 bp in length  
\* 21883 21982: gap of unknown length  
\* 21983 25547: contig of 3565 bp in length  
\* 25548 25647: gap of unknown length  
\* 25648 28469: contig of 2822 bp in length  
\* 28470 28569: gap of unknown length  
\* 28570 32543: contig of 3974 bp in length  
\* 32544 32643: gap of unknown length  
\* 32644 39611: contig of 6968 bp in length

```
* 39612 39711: gap of unknown length
* 39712 48242: contig of 8531 bp in length
* 48243 48342: gap of unknown length
* 48343 54805: contig of 6463 bp in length
* 54806 62691: contig of 7786 bp in length
* 62692 62791: gap of unknown length
* 62792 70784: contig of 7993 bp in length
* 70785 70884: gap of unknown length
* 70885 80151: contig of 9267 bp in length
* 80152 80252: gap of unknown length
* 80253 89912: contig of 9661 bp in length
* 89913 99013: gap of unknown length
* 99014 99541: contig of 9429 bp in length
* 99542 114999: gap of unknown length
* 115000 129243: contig of 15458 bp in length
* 129244 129343: contig of 14144 bp in length
* 129344 146645: gap of unknown length
* 146646 167157: contig of 17302 bp in length
* 167158 167258: gap of unknown length
* 167259 186353: contig of 20412 bp in length
* 186354 186453: gap of unknown length
* 186454 188860: contig of 19095 bp in length
* 188861 188961: contig of 2408 bp in length
* 188962 191426: gap of unknown length
* 191427 191426: contig of 2466 bp in length.
Location/Qualifiers
1. .191426
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-540N15"
1. .1754
/note="assembly_name:Contig10"
1855. .4193
/note="assembly_name:Contig11"
4294. .6143
/note="assembly_name:Contig12"
6244. .7809
/note="assembly_name:Contig13"
7910. .9758
/note="assembly_name:Contig14"
9859. .15360
/note="assembly_name:Contig20
clone end:T7
vector_side:left"
15461. .18198
/note="assembly_name:Contig15"
18299. .21882
/note="assembly_name:Contig16"
21983. .25547
/note="assembly_name:Contig17"
25648. .28469
/note="assembly_name:Contig18"
28570. .32543
/note="assembly_name:Contig19"
32644. .39611
/note="assembly_name:Contig21"
39712. .48242
/note="assembly_name:Contig22"
48343. .54805
/note="assembly_name:Contig23"
54906. .62691
/note="assembly_name:Contig24"
62792. .70784
/note="assembly_name:Contig25"
70885. .80151
/note="assembly_name:Contig26"
80252. .89912
/note="assembly_name:Contig27"

misc_feature
/note="assembly_name:Contig28"
99542. .114999
/note="assembly_name:Contig29"
115100. .129243
/note="assembly_name:Contig30"
129344. .146645
/note="assembly_name:Contig31"
146746. .167157
/note="assembly_name:Contig32"
167258. .186352
/note="assembly_name:Contig33"
186453. .188860
/note="assembly_name:Contig34"
188961. .191426
/note="assembly_name:Contig35"

ORIGIN
Query Match 87.0%; Score 17.4; DB 2; Length 191426;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAGTTTCAGTCAAGGATT 20
|||||
Db 156067 GAAGTTTCAGTCAAGGATT 156049

RESULT 12
AL627104/c
LOCUS
DEFINITION
AL627104 Mouse DNA sequence from clone RP23-412B2 on chromosome 4, complete
sequence.
ACCESSION AL627104
VERSION AL627104.16 GI:22204272
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.; EMBL; Sw.;
SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
```



```

repeat_region 6636. .6747 /rpt_family="MER112"
repeat_region complement(6825. .6919)
repeat_region 7030. .7113 /rpt_family="MIR"
repeat_region 7284. .7316 /rpt_family="MIR"
repeat_region complement(7317. .7612)
repeat_region /rpt_family="AluSg"
repeat_region 7700. .7738
repeat_region /rpt_family="AT_rich"
repeat_region 7848. .7934
repeat_region /rpt_family="MIR"
repeat_region complement(7975. .8079)
repeat_region /rpt_family="MER112"
repeat_region 8080. .8383
repeat_region /rpt_family="AluSx"
repeat_region complement(8384. .8422)
repeat_region /rpt_family="MER112"
repeat_region 8456. .8490
repeat_region /rpt_family="AT_rich"
repeat_region complement(8617. .8770)
repeat_region /rpt_family="FRAM"
repeat_region 9364. .9488
repeat_region /rpt_family="L2"
repeat_region 10703. .11425
repeat_region /rpt_family="L1MED"
repeat_region 11599. .11897
repeat_region /rpt_family="L1MED"
repeat_region 12328. .12414
repeat_region /rpt_family="L1MED"
repeat_region complement(12445. .12582)
repeat_region /rpt_family="MLT1F1"
repeat_region 12592. .12650
repeat_region /rpt_family="MIR3"
repeat_region 12990. .13101
repeat_region /rpt_family="MIR"
repeat_region complement(13144. .13350)
repeat_region /rpt_family="L2"
repeat_region 13351. .13511
repeat_region /rpt_family="MERSA"
repeat_region complement(13512. .13841)
repeat_region /rpt_family="L2"
repeat_region 15177. .15587
repeat_region /rpt_family="MLT1C"
repeat_region 16117. .16147
repeat_region /rpt_family="TTTG)n"
repeat_region complement(16150. .16436)
repeat_region /rpt_family="AluJo"
repeat_region complement(16438. .16572)
repeat_region /rpt_family="MIR"
repeat_region 16852. .17434
repeat_region /rpt_family="LTR50"
repeat_region complement(17434. .17813)
repeat_region /rpt_family="MLT11"
repeat_region complement(17855. .18179)
repeat_region /rpt_family="L2"
repeat_region complement(18474. .19179)
repeat_region /rpt_family="L1MB3"
repeat_region 19585. .19670
repeat_region /rpt_family="L2"
repeat_region complement(19698. .19725)
repeat_region /rpt_family="MIR"
repeat_region 20513. .20679
repeat_region /rpt_family="FRAM"
repeat_region 20804. .20900
repeat_region /rpt_family="MIR"
repeat_region 20901. .21259
repeat_region /rpt_family="THE1B"
repeat_region 21260. .21279
repeat_region /rpt_family="MIR"
repeat_region complement(21320. .21428)
/rpt_family="MER20"
complement(21653. .21891)
/rpt_family="MIR"

Query Match 87.0%; Score 17.4; DB 9; Length 194546;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GAAGTTGAGTCAAGGATT 20
|||||
Db 134063 GAAGCTTGAGTCAAGGATT 134081

RESULT 14
AC125890 208945 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-281J17, *** SEQUENCING IN PROGRESS
DEFINITION ***
AC125890
AC125890.3 GI:25092392
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 208945)
REFERENCE
AUTHORS
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gregoregis,B., Geer,K., Gili,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpthy,S., Kelly,S., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naif,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Snead,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,I., Thomas,N., Thomas,S., Tingey,A., Tjebk,B., Umanil,K.,
Valas,R., Vera,V., Villanar,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

```

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.  
Direct Submission  
Unpublished  
2 (bases 1 to 208945)  
Worley, K. C.  
Direct Submission  
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 208945)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 19, 2002 this sequence version replaced gi:23195060.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GWLQ  
Center clone name: CH230-281J17  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 167086 bases at least Q40  
Consensus quality: 170691 bases at least Q30  
Consensus quality: 173471 bases at least Q20  
Estimated insert size: 176890; sum-of-contigs estimation  
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 208945: contig of 208945 bp in length.

## FEATURES

source  
1. .208945  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-281J17"  
misc\_feature  
1. .1485  
/notes="wgs\_end\_extension  
clone end:T7"  
2244..3125  
misc\_feature  
/note="clone\_boundary  
clone end:T7"  
site:  
end\_sequence:B2234356"  
misc\_feature  
8551..10700

misc\_feature  
/note="wgs\_contig"  
24029..26647  
/note="wgs\_contig"  
90978..93358  
/note="wgs\_contig"  
140618..141928  
/note="wgs\_contig"  
207618..208514  
/note="clone\_boundary  
clone\_end:Sp6  
site:  
end\_sequence:B2234357"

ORIGIN  
Query Match 87.0%; Score 17.4; DB 2; Length 208945;  
Best Local Similarity 94.7%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGAAGGTTGAGTCAAGGAT 19  
||||| |||||||||  
Db 167140 GGAAGGTTGAGTCAAGGAT 167158  
RESULT 15  
AC095610/c  
LOCUS  
Rattus norvegicus clone CH230-7M13, WORKING DRAFT SEQUENCE.  
AC095610  
AC095610.5 GI:30467191  
HTG; HTGS PHASE2; HTGS DRAFT; HTGS\_FULLTOP.  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 224855)  
AUTHORS  
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, W.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladik, S. L., Hodgson, A., Hoggues, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensu, H., Loulseg, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwankweme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puaio, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

TITLE		TITLE
JOURNAL		AUTHORS
REFERENCE		TITLE
AUTHORS		JOURNAL
TITLE		
JOURNAL		

## FEATURES

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 15:09:47 ; Search time 13.2551 Seconds  
(without alignments)  
8932.047 Million cell updates/sec

Title: US-09-899-276C-8

Perfect score: 20

Sequence: 1 ggaaggttgagcaaggatt 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_15Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	ADH13945 Human MCP
2	17	85.0	21	6	ADH13947 Human mon
3	17	85.0	600	6	ADH13938 Human mon
4	17	85.0	11793	12	ADO03803 Human Cc1
5	16.8	84.0	1152	5	AAS69735 DNA encod
6	16.8	84.0	1604	3	AAC42300 Arabidops
7	16.8	84.0	2778	3	AAC48378 Arabidops
8	16.8	84.0	59002	12	ADQ97989 Mouse can
9	16.4	82.0	939	3	AAC57918 Arachidon
10	16.4	82.0	110000	12	ADQ97050_5 Continuation (6 of
11	16	80.0	84495	6	AAS20588 Human met
12	15.8	79.0	399	4	AAL20747 Human bre
13	15.8	79.0	433	4	AAL11857 Human bre
14	15.8	79.0	459	11	ACN82053 Breast ca
15	15.8	79.0	462	4	AAS56196 Human imm
16	15.8	79.0	533	4	AAL16869 Human cDN
17	15.8	79.0	758	11	ACN90528 Breast ca
18	15.8	79.0	1064	10	ADG90903 Hepatic s
19	15.8	79.0	1324	10	ADG90914 Hepatic s
20	15.8	79.0	1394	10	ADG90913 Hepatic s

c 21	15.8	79.0	1763	13	ADS10115	AdS10115 Human the
c 22	15.8	79.0	2582	3	AAC54273	Aac54273 Arabidops
c 23	15.8	79.0	2657	10	ADG90912	Adg90912 Hepatic s
c 24	15.8	79.0	2699	4	AAL17761	Aal17761 Human cDN
c 25	15.8	79.0	2846	10	ADG90904	Adg90904 Hepatic s
c 26	15.8	79.0	2946	10	ADG90901	Adg90901 Hepatic s
c 27	15.8	79.0	3818	10	ADG90900	Adg90900 Hepatic s
c 28	15.8	79.0	4594	10	ADG90887	Adg90887 Hepatic s
c 29	15.8	79.0	5186	10	ADG90899	Adg90899 Hepatic s
c 30	15.8	79.0	5753	10	ADG90911	Adg90911 Hepatic s
c 31	15.8	79.0	6748	10	ADG90891	Adg90891 Hepatic s
c 32	15.8	79.0	7036	10	ADG90909	Adg90909 Hepatic s
c 33	15.8	79.0	7037	10	ADG90907	Adg90907 Hepatic s
c 34	15.8	79.0	7065	10	ADG90908	Adg90908 Hepatic s
c 35	15.8	79.0	7096	10	ADG90889	Adg90889 Hepatic s
c 36	15.8	79.0	7101	10	ADG90910	Adg90910 Hepatic s
c 37	15.8	79.0	7144	10	ADG90895	Adg90895 Hepatic s
c 38	15.8	79.0	7161	10	ADG90894	Adg90894 Hepatic s
c 39	15.8	79.0	7237	10	ADG90905	Adg90905 Hepatic s
c 40	15.8	79.0	8200	10	ADG90898	Adg90898 Hepatic s
c 41	15.8	79.0	8570	10	ADG90893	Adg90893 Hepatic s
c 42	15.8	79.0	9260	8	ABX94633	Abx94633 C. elegans
c 43	15.8	79.0	12309	4	AAK89483	Aak89483 Human dig
c 44	15.8	79.0	12309	4	AAK90834	Aak90834 Human dig
c 45	15.8	79.0	96587	9	ADA02984	Ada02984 Human MAP

ALIGNMENTS

RESULT 1

ADH13945

ID ADH13945 standard; DNA; 20 BP.

XX

AC ADH13945;

XX 11-MAR-2004 (first entry)

XX Human MCP-1 related oligonucleotide SEQ ID NO:8.

ss; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
cervical carcinoma.

OS Homo sapiens.

XX

PN EP1170372-A1.

XX

PD 09-JAN-2002.

XX 06-JUL-2000; 2000BP-00114560.

PF 06-JUL-2000; 2000BP-00114560.

PR 06-JUL-2000; 2000BP-00114560.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Roosl F., Soto U., Coy J., Finzer P., Delius H., Poustka A;

XX Zur Hausen H., Patzelt A;

XX WPI; 2002-165895/22.

Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
useful in gene therapy, for treating atherosclerosis and cancer.

PS Disclosure; SEQ ID NO 8; 30pp; English.

XX The invention relates to a novel nucleic acid molecule (I) comprising a  
sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
having the biological activity of MCP-1. A protein encoded by a nucleic  
acid of the invention has cytostatic, and antiarteriosclerotic activity.  
XX A nucleic acid of the invention may have a use in gene therapy. A  
compound of the invention is useful in the preparation of a medicament  
for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A

Tue Aug 9 17:01:29 2005

us-09-899-276c-8.rng

CC pharmaceutical composition of the invention is useful for the treatment  
 CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
 CC atherosclerosis or cancer. The present sequence is used in the  
 CC exemplification of the invention.

XX SQ Sequence 20 BP; 6 A; 1 C; 8 G; 5 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 20; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20  
 |||||  
 Db 1 GGAAGGTTGAGTCAAGGATT 20

RESULT 2  
 ADH13947  
 ID ADH13947 standard; DNA; 21 BP.  
 XX AC ADH13947;  
 XX DT 11-MAR-2004 (first entry)  
 XX Human monocyte-chemoattractant-protein-1 AP-1 binding site SEQ ID NO:10.  
 DE human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
 KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
 KW cervical carcinoma; ss.  
 XX OS Homo sapiens.

XX EP1170372-A1.  
 XX 09-JAN-2002.  
 XX 06-JUL-2000; 2000EP-00114560.  
 XX 06-JUL-2000; 2000EP-00114560.  
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX Roessl F, Soto U, Coy J, Finzer P, Delius H, Poustka A;  
 XX Zur Hausen H, Patzelt A;  
 XX WPI; 2002-165895/22.  
 XX Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
 XX useful in gene therapy, for treating atherosclerosis and cancer.

XX Example 1; SEQ ID NO 10; 30pp; English.  
 XX The invention relates to a novel nucleic acid molecule (I) comprising a  
 CC sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
 CC having the biological activity of MCP-1. A protein encoded by a nucleic  
 CC acid of the invention has cytostatic, and antiarteriosclerotic activity.  
 CC A nucleic acid of the invention may have a use in gene therapy. A  
 CC compound of the invention is useful in the preparation of a medicament  
 CC for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A  
 CC pharmaceutical composition of the invention is useful for the treatment  
 CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
 CC atherosclerosis or cancer. The present sequence is used in the  
 CC exemplification of the invention.

XX SQ Sequence 21 BP; 6 A; 1 C; 9 G; 5 T; 0 U; 0 Other;  
 Query Match 85.0%; Score 17; DB 6; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGG 17  
 |||||  
 Db 1 GGAAGGTTGAGTCAAGG 17

RESULT 3  
 ADH13938  
 ID ADH13938 standard; DNA; 600 BP.  
 XX AC ADH13938;  
 XX DT 11-MAR-2004 (first entry)  
 XX Human monocyte-chemoattractant-protein-1 related DNA SEQ ID NO:1.  
 DE ds; human; monocyte-chemoattractant-protein-1; MCP-1; cytostatic;  
 KW antiarteriosclerotic; gene therapy; atherosclerosis; cancer;  
 KW cervical carcinoma.

XX OS Homo sapiens.  
 XX EP1170372-A1.  
 XX 09-JAN-2002.  
 XX 06-JUL-2000; 2000EP-00114560.  
 XX 06-JUL-2000; 2000EP-00114560.  
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX Roessl F, Soto U, Coy J, Finzer P, Delius H, Poustka A;  
 XX Zur Hausen H, Patzelt A;  
 XX WPI; 2002-165895/22.  
 XX Novel nucleic acid molecule encoding monocyte-chemoattractant-protein-1,  
 XX useful in gene therapy, for treating atherosclerosis and cancer.  
 XX Disclosure; SEQ ID NO 1; 30pp; English.

XX The invention relates to a novel nucleic acid molecule (I) comprising a  
 CC sequence encoding monocyte-chemoattractant-protein-1 (MCP-1) or a protein  
 CC having the biological activity of MCP-1. A protein encoded by a nucleic  
 CC acid of the invention has cytostatic, and antiarteriosclerotic activity.  
 CC A nucleic acid of the invention may have a use in gene therapy. A  
 CC compound of the invention is useful in the preparation of a medicament  
 CC for the treatment of atherosclerosis or cancer e.g. cervical carcinoma. A  
 CC pharmaceutical composition of the invention is useful for the treatment  
 CC of diseases associated with dysregulation of MCP-1 expression, e.g.  
 CC atherosclerosis or cancer. The present sequence is used in the  
 CC exemplification of the invention.

XX SQ Sequence 600 BP; 213 A; 101 C; 150 G; 136 T; 0 U; 0 Other;  
 Query Match 85.0%; Score 17; DB 6; Length 600;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGG 17  
 |||||  
 Db 158 GGAAGGTTGAGTCAAGG 174

RESULT 4  
 ADO03803  
 ID ADO03803 standard; DNA; 11793 BP.  
 XX AC ADO03803;  
 XX DT 12-AUG-2004 (first entry)  
 XX Human Ccl2 gene and enhancer region DNA SeqID 4.  
 DE human; ds; animal model; age-related macular degeneration; AMD;  
 KW gene knockout; Ccl2-deficient; Ccr2-deficient; drusen;

KW lipofuscin accumulation; Bruch's membrane; retinal degeneration;  
KW choroidal neovascularisation; ophthalmological; gene therapy.

XX Homo sapiens.

XX WO2004041160-A2.

XX 21-MAY-2004.

XX 16-OCT-2003; 2003WO-US032933.

XX 30-OCT-2002; 2002US-0422096P.

XX (KENT ) UNIV KENTUCKY RES FOUND.

XX Ambati J;

XX WPI; 2004-400512/37.

XX Testing candidate drug for treating age-related macular degeneration, by  
PT administering drug to Ccl2-deficient, Ccr2-deficient knockout mouse, and  
PT analyzing development or regression of drusen and/or lipofuscin  
PT accumulation in eye.

XX Disclosure; SEQ ID NO 4; 64pp; English.

XX This invention relates to a novel methods and animal models for testing  
CC candidate drugs that can be used for the treatment or prevention of age-  
CC related macular degeneration (AMD). Specifically, it refers to  
CC administering a candidate drug to gene knockout mice, in particular Ccl2-  
CC deficient, Ccr2-deficient and/ or a Ccl2-deficient/Ccr2-deficient dual  
CC knockout mouse. The present invention describes analysing the knockout  
CC mouse eye for development or regression of drusen and/ or lipofuscin  
CC accumulation, as well as for the effect of a candidate drug on Bruch's  
CC membrane, retinal degeneration and/ or choroidal neovascularisation.  
CC Accordingly, such compositions exhibit ophthalmological activities and  
CC can be used for gene therapy purposes. This polynucleotide sequence is  
CC the human Ccl2 gene and enhancer region DNA of the invention.

SQ Sequence 11793 BP; 3209 A; 2913 C; 2559 G; 3112 T; 0 U; 0 Other;

Query Match 85.0%; Score 17; DB 12; Length 11793;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGG 17

Db 10136 GGAAGGTTGAGTCAAGG 10152

RESULT 5

AAS69735/c

ID AAS69735 standard; cDNA; 1152 BP.

XX AAS69735;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5539.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG05548.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 5539; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 1152 BP; 318 A; 286 C; 253 G; 295 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 5; Length 1152;

Best Local Similarity 90.0%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20

Db 843 GGAAGATTCAAGTCAAGGATT 824

RESULT 6

AAC42300

ID AAC42300 standard; DNA; 1604 BP.

XX AAC42300;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 35028.

XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0134449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 14-MAY-1999; 99US-0134768P.  
PR 18-MAY-1999; 99US-0134941P.  
PR 19-MAY-1999; 99US-0135124P.  
PR 20-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 24-MAY-1999; 99US-0136021P.  
PR 25-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 07-JUN-1999; 99US-0138094P.  
PR 08-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 14-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139492P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139899P.  
PR 22-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140695P.  
PR 23-JUN-1999; 99US-0140823P.  
PR 24-JUN-1999; 99US-0140931P.  
PR 28-JUN-1999; 99US-0141287P.  
PR 29-JUN-1999; 99US-0141842P.  
PR 30-JUN-1999; 99US-0142154P.  
PR 01-JUL-1999; 99US-0142055P.  
PR 02-JUL-1999; 99US-0142390P.  
PR 06-JUL-1999; 99US-0142803P.  
PR 08-JUL-1999; 99US-0142920P.  
PR 09-JUL-1999; 99US-0142977P.  
PR 12-JUL-1999; 99US-0143542P.  
PR 13-JUL-1999; 99US-0143624P.  
PR 14-JUL-1999; 99US-0144005P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.



PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160880P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	28-SEP-1999;	99US-0155659P.			
PR	29-SEP-1999;	99US-0155598P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157533P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
Query Match 84.0%; Score 16.8; DB 3; Length 2778;					
Best Local Similarity 90.0%; Pred. No. 1.2e+02;					
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	GGAGGTTGAGTCAAGGATT 20			
Db	2099	GGAGGTTGAAGCAAGGATT 2118			
RESULT 8					
ID	ADQ97989/c				
ID	ADQ97989 standard; DNA; 59002 BP.				
XX	ADQ97989;				
XX	07-OCT-2004 (first entry)				
XX	Mouse cancer associated sequence MD11-037, SEQ ID 966.				
XX	Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.				
XX	Mus musculus.				
XX	WO2004060304-A2.				
XX	22-JUL-2004.				
XX	22-DEC-2003; 2003WO-US041389.				
XX	27-DEC-2002; 2002US-00330773.				
XX	(SAGR-) SAGRES DISCOVERY INC.				
XX	Morris DW, Malandro MS;				
XX	WPI; 2004-543781/52.				
XX	New isolated cancer associated nucleic acids comprising at least 10				
XX	contiguous nucleotides, useful for diagnosing, preventing and/or treating				
XX	cancers such as leukemia and lymphoma.				
XX	Claim 1; SEQ ID NO 966; 199pp; English.				
XX	The present invention relates to cancer associated sequences (ADQ97025-				
XX	ADQ98004). The sequences are useful for the diagnosis, prevention and/or				
XX	treatment of cancer, such as leukemia and lymphoma. Note: The sequence				
XX	data for this patent did not form part of the printed specification, but				
XX	was obtained in electronic formate directly from WIPO at				

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 59002 BP; 16030 A; 11777 C; 12009 G; 18102 T; 0 U; 1084 Other;  
SQ Best Local Similarity 84.0%; Score 16.8; DB 12; Length 59002;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20  
||||| ||||||| |||||  
DB 7393 GGAAGGATGAGTCAAGCATT 7374

RESULT 9  
AAC57918  
ID AAC57918 standard; DNA; 939 BP.

XX AAC57918;

DT 25-JAN-2001 (first entry)

XX Arachidonic acid metabolism related genomic biallelic marker #552.

XX Human; biallelic marker; arachidonic acid metabolism; genotyping;  
KW detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;  
KW single nucleotide polymorphism; hybridisation assay; sequencing assay;  
KW specific amplification assay; identification; ERM; 12-LO-REM;  
KW eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.

XX Homo sapiens.

XX WO200047771-A2.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-IB000184.

XX 12-FEB-1999; 99US-0119917P.

XX 23-MAR-1999; 99US-00275267.

XX 07-MAY-1999; 99US-0133200P.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I;

XX WPI; 2000-571881/53.

XX Novel biallelic markers useful for detecting conditions and genotypes  
PT associated with arachidonic acid metabolism.

XX Claim 13; Page 693; 802pp; English.

XX The present invention describes polynucleotides including biallelic  
CC markers derived from genes involved in arachidonic acid metabolism and  
CC from genomic regions flanking those genes. Methods from the present  
CC invention may be used to select individuals for clinical trials and  
CC predict responses to treatment with drugs. The polynucleotides may be  
CC used in hybridisation assays, sequencing assays and specific  
CC amplification assays for identifying an eicosanoid-related biallelic  
CC marker (EREM) or 12-LO-related biallelic marker, and for amplifying a  
CC segment of nucleotides containing an ERM. The polynucleotides are useful  
CC in diagnostic kits. The markers may be used to detect conditions and  
CC genotypes associated with arachidonic acid metabolism. AAC57367 to  
CC AAC58018 and AAB24019 and AAB24020 represent sequences used in the  
CC exemplification of the present invention. N.B. Polymorphic bases (single  
CC nucleotide polymorphisms also known as SNPs) in the polynucleotide  
CC sequences from the present invention have been given as their  
CC corresponding degenerate bases e.g. a polymorphic base of C or T has been  
CC given as Y

XX Sequence 939 BP; 249 A; 187 C; 246 G; 256 T; 0 U; 1 Other;

SQ Query Match 82.0%; Score 16.4; DB 3; Length 939;

Best Local Similarity 94.4%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGA 18  
||||| ||||||| |||||||  
DB 548 GGAAGGTTGAGTCAAGCA 565

RESULT 10  
ADQ97050\_5

Continuation (6 of 7) of ADQ97050 from base 500001 (Human cancer associated sequence HD  
WP Sequence split into 7 fragments LOCUS ADQ97050 Accession Adq97050

WP	Fragment Name	Begin	End
WP	ADQ97050_0	1	110000
WP	ADQ97050_1	100001	210000
WP	ADQ97050_2	200001	310000
WP	ADQ97050_3	300001	410000
WP	ADQ97050_4	400001	510000
WP	ADQ97050_5	500001	610000
WP	ADQ97050_6	600001	687411

Query Match 82.0%; Score 16.4; DB 12; Length 110000;  
Best Local Similarity 94.4%; Pred. No. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAGGTTGAGTCAAGGAT 19  
||||| ||||||| |||||||  
DB 66453 GAAGGTTGATTCAGGAT 66470

RESULT 11

AAS20588/c  
ID AAS20588 standard; DNA; 84495 BP.

XX AAS20588;

XX 23-APR-2002 (first entry)

XX Human methionine aminopeptidase protease genomic DNA.

XX Human; methionine aminopeptidase; protease; adrenal cortico adenoma; ds;  
KW Cushing's syndrome; prostate embryonal carcinoma; colon tumour; brain;  
KW hepatocellular carcinoma; foetal lung; testis; b cell; kidney; prostate;  
KW gene.

XX Homo sapiens.

XX US6329188-B1.

XX 11-DEC-2001.

XX 05-MAR-2001; 2001US-00797906.

XX 02-MAR-2001; 2001US-00797000.

XX (PEKE ) PE CORP NY.

XX Yan X, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-121135/16.

XX Novel nucleic acid molecule encoding methionine aminopeptidase, useful  
PT for the development of human therapeutics and diagnostic compositions.

XX Claim 1; Fig 3; 109pp; English.

XX The invention relates to a human methionine aminopeptidase protease  
CC polypeptide and its associated DNA sequence. The sequences are useful for  
CC producing protease proteins including methionine aminopeptidase and as  
CC targets for the development of human therapeutic agents that modulate  
CC protease activity in cells and tissues that express the polypeptide. The  
CC sequences of the invention are expressed in adrenal cortico adenoma for  
CC Cushing's syndrome, prostate embryonal carcinoma, colon tumour, adult

CC brain, hepatocellular carcinoma, foetal lung, testis and b cell, kidney  
 CC and prostate. This sequence represents genomic DNA encoding the human  
 CC methionine aminopeptidase of the invention  
 XX  
 SQ Sequence 84495 BP; 23805 A; 16441 C; 17154 G; 25835 T; 0 U; 1260 Other;

Query Match 80.0%; Score 16; DB 6; Length 84495;  
 Best Local Similarity 100.0%; Pred. No. 5e+02;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAG 16  
 |||||  
 Db 34427 GGAAGGTTGAGTCAAG 34412

RESULT 12  
 AAL20747  
 ID AAL20747 standard; cDNA; 399 BP.  
 XX  
 AC AAL20747;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 DE Human breast cancer expressed polynucleotide 13204.  
 XX  
 KW Human; breast cancer; cell marker; cytostatic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200151628-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 10-JAN-2001; 2001WO-US000798.  
 XX  
 PR 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0192099P.  
 PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 XX Lillie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 XX WPI; 2001-451856/48.  
 XX  
 PF New peptide useful as a marker for the diagnosis of breast cancer.  
 XX  
 PS Claim 1; Page 2342; 3695pp; English.  
 XX  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 XX  
 SQ Sequence 399 BP; 123 A; 55 C; 132 G; 89 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 399;  
 Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGTTGAGTCAAGATT 20  
 |||||  
 Db 200 GGAAGTTGAGTCAAGATT 218

RESULT 13  
 AAL11857  
 ID AAL11857 standard; cDNA; 433 BP.  
 XX  
 AC AAL11857;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 DE Human breast cancer expressed polynucleotide 4314.  
 XX  
 KW Human; breast cancer; cell marker; cytostatic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200151628-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 10-JAN-2001; 2001WO-US000798.  
 XX  
 PR 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0192099P.  
 PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 XX Lillie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 XX WPI; 2001-451856/48.  
 XX  
 PF New peptide useful as a marker for the diagnosis of breast cancer.  
 XX  
 PS Claim 1; Page 783; 3695pp; English.  
 XX  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 XX  
 SQ Sequence 433 BP; 122 A; 69 C; 143 G; 97 T; 0 U; 2 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 433;  
 Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAAGTTGAGTCAAGATT 20  
 |||||  
 Db 240 GGAAGTTGAGTCAAGATT 258

RESULT 14  
 ACN82053  
 ID ACN82053 standard; DNA; 459 BP.  
 XX  
 AC ACN82053;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Breast cancer related marker, seq id 3203.  
 XX  
 KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.  
 XX  
 OS Homo sapiens.

```
XX  US2003099974-A1.
XX  29-MAY-2003.
XX  18-JUL-2002; 2002US-00198846.
XX  18-JUL-2001; 2001US-0306220P.
XX  (MILL-) MILLENNIUM PHARM INC.
XX  Lillie J, Xu Y, Wang Y, Steinmann K;
XX  WPI; 2003-787014/74.
XX  Novel isolated polypeptide associated with breast cancer, useful for
XX  detecting presence of polypeptide in sample, as a marker for breast
XX  cancer.
XX  Disclosure; SEQ ID NO 3203; 36pp; English.
XX  The invention relates to an isolated polypeptide (I) associated with
XX  breast cancer which is encoded by a nucleic acid molecule comprising a
XX  nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX  the polypeptide of the invention. The activity of the polypeptide of the
XX  invention may be described as cytostatic. The antibody is useful for
XX  detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX  invention are useful in the detection of breast tumours. (I) is useful as
XX  a marker for breast cancer and in breast cancer therapy. Sequences given
XX  in records ACN78851-ACN92934 represent nucleic acid markers associated
XX  with breast cancer. Note: The sequence listing does not form part of the
XX  specification but may be obtained in electronic format from the USPTO web
XX  site at seqdata.uspto.gov/sequence.html?docID=20030099974
XX  SQ Sequence 459 BP; 130 A; 73 C; 148 G; 103 T; 0 U; 5 Other;

Query Match 79.0%; Score 15.8; DB 11; Length 459;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGGTTGAGTCAGGATT 20
DB 249 GAAGGTGGAGTCAGGATT 267

RESULT 15
AAK56196/c
ID AAK56196 standard; cDNA; 462 BP.
XX
AC AAK56196;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1256.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
```

CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 462 BP; 96 A; 134 C; 137 G; 91 T; 0 U; 4 Other;  
Query Match 79.0%; Score 15.8; DB 4; Length 462;  
Best Local Similarity 89.5%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 GGAAGGTTGAGTCAAGGAT 19  
Db 183 GCAAGGTTGAGTTAAGGAT 165  
Search completed: August 4, 2005, 08:31:32  
Job time : 16.2551 secs

20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 20-OCT-2000; 2000US-0244617P.  
PR 01-NOV-2000; 2000US-0244647P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249224P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 06-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX P-PSDB; AAK83415.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Claim 1; SEQ ID NO 1256; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
XX amino acid sequences given in AAK91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 04:04:35 ; Search time 94.3968 Seconds  
(without alignments)  
8064.741 Million cell updates/sec

Title: US-09-899-276C-8

Perfect score: 20

Sequence: 1 ggaaggttgagtcaggatt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	18	90.0	400	1	AJ516664
2	17.4	87.0	361	9	CR008527
3	17.4	87.0	453	8	AZ795398
4	17.4	87.0	580	1	AI100780
5	17.4	87.0	700	9	AG180768
6	17.4	87.0	832	8	AQ998174
7	17	85.0	284	9	CC836847
8	17	85.0	537	8	AZ380590
9	17	85.0	771	4	BG721914
10	16.8	84.0	281	9	CC558114
11	16.8	84.0	305	9	CC675901
12	16.8	84.0	308	1	AJ792516
13	16.8	84.0	335	9	EX160484
14	16.8	84.0	345	6	BY663202
15	16.8	84.0	351	2	AW321997
16	16.8	84.0	352	9	CG551634
17	16.8	84.0	363	9	CG658504
18	16.8	84.0	365	1	AW098203
19	16.8	84.0	476	2	AJ794272
20	16.8	84.0	514	9	CE463871
21	16.8	84.0	521	8	CC424518
22	16.8	84.0	525	2	AW228043
23	16.8	84.0	550	9	CG243714
24	16.8	84.0	560	6	CD799807

C 25	16.8	84.0	575	8	BZ820561	BZ820561
C 26	16.8	84.0	582	5	BP259697	BP259697
C 27	16.8	84.0	589	8	AQ299288	HS 2231.A
C 28	16.8	84.0	615	5	BUT91070	60414287
C 29	16.8	84.0	619	8	AZ032895	AZ032895
C 30	16.8	84.0	627	5	BQ403091	GA_Ea005
C 31	16.8	84.0	634	4	BG442583	GA_Ea001
C 32	16.8	84.0	637	5	BW319959	BW319959
C 33	16.8	84.0	653	2	BB458766	BB458766
C 34	16.8	84.0	656	8	AZ790649	2M0039E16
C 35	16.8	84.0	668	2	AW037663	AW037663
C 36	16.8	84.0	668	7	CN465627	9853.1 AE
C 37	16.8	84.0	670	5	BW327210	BW327210
C 38	16.8	84.0	671	5	BQ587296	BQ587296
C 39	16.8	84.0	671	5	BQ587698	E012340w-
C 40	16.8	84.0	675	5	BQ412080	BQ412080
C 41	16.8	84.0	688	4	BG441630	GA_Ea001
C 42	16.8	84.0	699	6	CD823901	CD823901
C 43	16.8	84.0	718	6	CB597302	CB597302
C 44	16.8	84.0	724	9	CL652087	PR10113d
C 45	16.8	84.0	737	6	CB593048	CB593048

## ALIGNMENTS

RESULT 1  
AJ516664  
LOCUS  
DEFINITION  
AJ516664 Mytilus galloprovincialis haemolymph, gills, digestive gland, foot, adductor muscles and mantle Mytilus galloprovincialis cDNA clone GPPD00678, mRNA sequence.  
ACCESSION  
AJ516664  
VERSION  
AJ516664.1 GI:37650241  
KEYWORDS  
EST:  
SOURCE  
Mytilus galloprovincialis (Mediterranean mussel)  
ORGANISM  
Mytilus galloprovincialis  
REFERENCE  
1 (bases 1 to 400)  
AUTHORS  
Venier,P., Pallavicini,A., De Nardi,B. and Lanfranchi,G.  
TOWARDS A CATALOGUE OF GENES TRANSCRIBED IN MULTIPLE TISSUES OF MYTILUS GALLOPROVINCIALIS  
JOURNAL  
Gene 314, 29-40 (2003)  
MEDLINE  
22890058  
PubMed  
14527715  
COMMENT  
Contact: Venier P  
Biology  
University of Padova  
Via Ugo Bassi 58/b, Italy, 34100, Italy.  
Location/Qualifiers  
1. 400  
/organism="Mytilus galloprovincialis"  
/mol\_type="mRNA"  
/db\_xref="taxon:29158"  
/clone="GPPD00678"  
/tissue type="haemolymph, gills, digestive gland, foot, adductor muscles and mantle"  
/clone lib="Mytilus galloprovincialis haemolymph, gills, digestive gland, foot, adductor muscles and mantle"

Query Match 90.0%; Score 18; DB 1; Length 400;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## ORIGIN

Qy 3 AAGGTGAGTCAAGGATT 20  
Db 168 AAGGTGAGTCAAGGATT 185

## RESULT 2

CR008527/c

LOCUS CR008527 361 bp DNA linear GSS 05-JUL-2004  
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN202111, genomic survey sequence.  
ACCESSION CR008527  
VERSION GSS; genome survey sequence; MICER.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER  
FEATURES  
source  
1..361  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN202111"  
/clone\_lib="MHPN"

ORIGIN  
Query Match 87.0%; Score 17.4; DB 9; Length 361;  
Best Local Similarity 94.7%; Pred. No. 6.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGAAGTTGAGTCAAGGAT 19  
|||||  
Db 139 GGAAGTTGAGGCAAGGAT 121

RESULT 3  
AZ795398 453 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0049B10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0049B10 R, genomic survey sequence.  
DEFINITION  
ACCESSION AZ795398  
VERSION GSS.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 453)  
AUTHORS Irlam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,K., Rose,K., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLUC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0049 row: B column: 10  
Seq primer: CACACAGGAACAGCATGACC  
Class: plasmid ends  
High quality sequence stop: 453.  
FEATURES  
source  
1..453  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"

/db\_xref="taxon:10090"  
/clone="UUGC2M0049B10"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
Query Match 87.0%; Score 17.4; DB 8; Length 453;  
Best Local Similarity 94.7%; Pred. No. 6.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGAAGTTGAGTCAAGGAT 19  
|||||  
Db 406 GGAAGTTGAGGCAAGGAT 424

RESULT 4  
AII100780/c 580 bp mRNA linear EST 31-JAN-1999  
LOCUS EST210069 Normalized rat brain, Bento Soares Rattus sp. cDNA clone RBBA18 3' end, mRNA sequence.  
DEFINITION  
ACCESSION AII100780  
VERSION AII100780.1 GI:3704516  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 580)  
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (REBT) Catalog & Rat Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
Seq primer: M13-21.  
FEATURES  
source  
1..580  
/organism="Rattus sp."  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):2023775"  
/db\_xref="taxon:10118"  
/clone="FBREA18"  
/clone\_lib="Normalized rat brain, Bento Soares"  
/note="Organ: Brain; Vector: pT713Fac; Site\_1: EcoRI; Site\_2: NotI"

ORIGIN



was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

# ORIGIN

Query Match 85.0%; Score 17; DB 8; Length 537;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGGTTGAGTCAAGG 17  
 Db 17 GGAAGGTTGAGTCAAGG 1

RESULT 9  
 LOCUS BG721914/c 771 bp mRNA linear EST 08-MAY-2001  
 DEFINITION 602695642P1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4827611 5', mRNA sequence.

ACCESSION BG721914  
 VERSION BG721914.1 GI:14001101  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 771)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgsbbs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10744 row: d column: 12  
 High quality sequence stop: 708.

# FEATURES

source  
 1..771  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4827611"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 97"  
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI, Site 2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH MGC Library."

# ORIGIN

Query Match 85.0%; Score 17; DB 4; Length 771;

Seq primer: SP6  
 Class: BAC ends  
 High quality sequence start: 97.  
 Location/Qualifiers  
 1..284  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBBC0482F11"  
 /lab\_host="E. coli DH10B"  
 /clone\_lib="ZMMBc"  
 /note="Vector: pTARBAC1.3; Site 1: BamHI; Site 2: BamHI"

# ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAGGTTGAGTCAAGG 17  
 Db 78 GGAAGGTTGAGTCAAGG 94

RESULT 8  
 LOCUS AZ380590/c 537 bp DNA linear GSS 02-OCT-2000  
 DEFINITION 1M0136P01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0136P01 R, genomic survey sequence.

ACCESSION AZ380590  
 VERSION AZ380590.1 GI:10494290  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 537)  
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Iellam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D. Weiss, R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: dunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0136 row: P column: 01  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 537.

# FEATURES

source  
 1..537  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC1M0136P01"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA

Best Local Similarity 100.0%; Pred. No. 1.2e+03; Indels 0; Gaps 0;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAGGTTGAGTCAGGATT 19  
 |||||

Db 698 AAGGTTGAGTCAGGATT 682

RESULT 10  
 CC558114  
 LOCUS 281 bp DNA linear GSS 19-JUN-2003  
 DEFINITION CH240\_466N17.T7 CHORI-240 Bos taurus genomic clone CH240\_466N17,  
 genomic survey sequence.  
 CC558114  
 ACCESSION CC558114.1 GI:31885477  
 VERSION  
 KEYWORDS  
 SOURCE Bos taurus (cow)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 281)  
 AUTHORS Holt,R., Stott,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,  
 Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,  
 Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,  
 Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,  
 Schein,J., Marra,M., de Jong,P., Keefe,J.W. and Kappes,S.M.  
 TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478  
 JOURNAL  
 COMMENT Unpublished (2003)  
 Other\_GSSs: CH240\_466N17.TARBAC13P2  
 Contact: Rob Holt  
 Sequencing  
 The British Columbia Cancer Agency Genome Science Centre  
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6  
 Tel: 604-877-6085  
 Fax: 604-877-6276  
 Email: rholt@bcgsc.ca  
 Clones are derived from the bovine BAC library CHORI-240  
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library  
 availability, please contact Pieter de Jong (pjejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/ordering/information.htm). This work  
 was undertaken as part of the International Bovine BAC Mapping  
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the  
 British Columbia Genome Sciences Centre, Canada.  
 Plate: 466 row: N column: 17  
 Seq primer: T7  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..281  
 /organism="Bos taurus"  
 /mol\_type="genomic DNA"  
 /strain="breed: Hereford"  
 /db\_xref="taxon:9913"  
 /clone="CH240\_466N17"  
 /sex="Male"  
 /cell\_type="Blood"  
 /clone\_lib="CHORI-240"  
 /note="Vector: PTARBAC1.3; Site\_1: MboI; Site\_2: MboI;  
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC  
 library (Male) produced by Pieter de Jong"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 281;  
 Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAGGATT 20  
 |||||

Db 215 GGAAGGTTGAGTCAGGATT 234

RESULT 11

CC675901/c  
 LOCUS 305 bp DNA linear GSS 19-JUN-2003  
 DEFINITION CGWEH38TH ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBma0558H04,  
 genomic survey sequence.  
 CC675901  
 ACCESSION CC675901.1 GI:32080599  
 VERSION  
 KEYWORDS  
 SOURCE Zea mays  
 ORGANISM  
 Zea mays

REFERENCE 1 (bases 1 to 305)

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
 TITLE Consortium for Maize Genomics  
 JOURNAL Unpublished (2002)  
 COMMENT Other\_GSSs: CGWEH38TV  
 Contact: Cathy Whitelaw  
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..305  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMMBma0558H04"  
 /clone\_lib="ZM 0.7 1.5\_KB"  
 /note="Vector: pBCSK; Site\_1: HincII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 305;  
 Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAGGATT 20  
 |||||

Db 89 GAAATGTTGAGTCAGGATT 70

RESULT 12  
 AJ792516  
 LOCUS 308 bp mRNA linear EST 11-AUG-2004  
 DEFINITION AJ792516 Antirrhinum majus whole plant Antirrhinum majus CDNA clone  
 018\_2\_11\_m02, mRNA sequence.  
 ACCESSION AJ792516  
 VERSION  
 KEYWORDS  
 SOURCE Antirrhinum majus (snapdragon)  
 ORGANISM  
 Antirrhinum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamiales; Plantaginales; Antirrhineae;  
 Antirrhinum.  
 REFERENCE 1 (bases 1 to 308)  
 AUTHORS Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.  
 TITLE Antirrhinum EST collection  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Schwarz-Sommer Z  
 Molekulare Pflanzen-genetik  
 MPI fuer Zuechtungs-forschung  
 Carl-von-Linne Weg 10, D-50829, Germany.  
 Location/Qualifiers  
 1..308  
 /organism="Antirrhinum majus"

```
/mol_type="mrna"
/db_xref="taxon:4151"
/clone="018.2.11.n02"
/tissue_type="whole plant"
/clone_lib="Antirrhinum majus whole plant"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 308;
Best Local Similarity 90.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20
    ||||| ||||| ||||| |||||
Db 134 GGAAGGTTTGTGTCACGGATT 153

RESULT 13
BX160484/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 335)
Humphray,S.J., Huckle,E. and Durham,J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 125C2. 125C2 is
part of the Dariokey BAC Library created by R. Plasterk and N.V.
keygene. Further details:
http://www.sanger.ac.uk/Projects/D_dario/.

FEATURES
source
1. .335
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-125C2"
/tissue_type="Testis"
/notes="vector pindigoBAC-536"

ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 335;
Best Local Similarity 90.0%; Pred. No. 1.3e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20
    ||||| ||||| ||||| |||||
Db 33 GGAAGGTTGAGTCAAGGATT 14

RESULT 14
BY663202/c
LOCUS
DEFINITION
Rathke's pouches Mus musculus cDNA clone K720019C17 3', mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 345)
Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,

Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Schonbach,C., Gojohori,T., Balderelli,R., Hill,D.P., Bult,C.,
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,
Ciochia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawai,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Sample,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
Shiraki,T., Tagami,M., Waki,K., Wataniki,A., Muramatsu,M. and
Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Michelle Brinkmeier and Sally Camper (
Dept. Human Genetics University of Michigan Medical School 4301
MSRB 3 1500 W. Medical Center Dr. Ann Arbor MI 48109-0638 USA )
whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
Location/Qualifiers
1. .345
/organism="Mus musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/clone="K720019C17"
/tissue_type="Rathke's pouches"

FEATURES
source
1. .345
/organism="Mus musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/clone="K720019C17"
/tissue_type="Rathke's pouches"
```

/dev\_stage="14.5 days embryo RP+/"  
/clone lib="RIKEN full-length enriched, 14.5 days embryo  
RP+/" Rathke's pouches"

Search completed: August 4, 2005, 14:32:52  
Job time : 102.397 secs

## ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 345;  
Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20  
|||||  
Db 154 GGAAGGTTGAGTCAAGGATT 135

## RESULT 15

AW321997/c 351 bp mRNA linear EST 25-JAN-2000  
u022c12.x1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:2609878 3'  
similar to gb:M38337 Mouse milk fat globule membrane protein E8  
mRNA, complete (MOUSE);, mRNA sequence.

ACCESSION AW321997.1 GI:6751541  
VERSION  
KEYWORDS  
SOURCE EST.

## ORGANISM

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 351)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

## JOURNAL

COMMENT  
Other ESTs: u022c12.Y1  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

MGI:1023738

Seq primer: -40UP from Gibco

High quality sequence stop: 223.

## FEATURES

source  
1..351  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:2609878"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam6"  
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 351;  
Best Local Similarity 90.0%; Pred. No. 1.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGATT 20  
|||||  
Db 307 GGAAGGTTGAGTCAAGGATT 288

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 07:01:49 ; Search time 4.12146 Seconds  
(without alignments)  
7940.282 Million cell updates/sec

Title: US-09-899-276C-8

Perfect score: 20

Sequence: 1 ggaaggttgagtcaggatt 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/PCUTUS COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/backfileseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	16.8	84.0	153642	4	US-09-949-016-12174
C 2	16.8	84.0	153643	4	US-09-949-016-15635
C 3	16.4	82.0	503	4	US-09-621-976-17595
C 4	16.4	82.0	939	3	US-09-641-638-552
C 5	16.4	82.0	939	4	US-10-170-097-552
C 6	16.4	82.0	670689	4	US-09-949-016-12505
C 7	16.4	82.0	670690	4	US-09-949-016-14207
C 8	16	80.0	84495	3	US-09-797-906-3
C 9	15.8	79.0	47555	4	US-09-949-016-16549
C 10	15.8	79.0	49673	4	US-09-949-016-12598
C 11	15.8	79.0	50563	4	US-09-949-016-15821
C 12	15.8	79.0	112132	3	US-09-741-150-3
C 13	15.8	79.0	112132	4	US-10-160-187-3
C 14	15.8	79.0	209210	4	US-09-949-016-15094
C 15	15.4	77.0	45229	4	US-09-949-016-12465
C 16	15.4	77.0	45300	4	US-09-949-016-13045
C 17	15.4	77.0	63930	4	US-09-949-016-12270
C 18	15.4	77.0	109690	4	US-09-949-016-13525
C 19	15.2	76.0	601	4	US-09-949-016-67753
C 20	15.2	76.0	601	4	US-09-949-016-79957
C 21	15.2	76.0	601	4	US-09-949-016-79958
C 22	15.2	76.0	601	4	US-09-949-016-136383
C 23	15.2	76.0	1923	4	US-09-134-000C-466
C 24	15.2	76.0	2004	4	US-09-214-982-2
C 25	15.2	76.0	2029	3	US-08-915-795-4
C 26	15.2	76.0	2029	4	US-09-296-275-4
C 27	15.2	76.0	2029	4	US-09-375-248-5

C 28	15.2	76.0	2117	4	US-09-949-016-5415	Sequence 5415, Ap
C 29	15.2	76.0	2469	3	US-08-858-207A-51	Sequence 51, Appl
C 30	15.2	76.0	2846	3	US-08-915-795-1	Sequence 1, Appli
C 31	15.2	76.0	2846	4	US-09-296-275-1	Sequence 1, Appli
C 32	15.2	76.0	12821	4	US-09-949-016-13589	Sequence 13589, A
C 33	15.2	76.0	18436	3	US-08-961-527-87	Sequence 87, Appl
C 34	15.2	76.0	24942	4	US-09-949-016-13738	Sequence 12738, A
C 35	15.2	76.0	28213	4	US-09-949-016-13738	Sequence 12738, A
C 36	15.2	76.0	28216	4	US-09-949-016-13652	Sequence 13652, A
C 37	15.2	76.0	42348	4	US-09-949-016-17157	Sequence 17157, A
C 38	15.2	76.0	44453	3	US-09-146-053-5	Sequence 5, Appli
C 39	15.2	76.0	56302	4	US-09-949-016-11892	Sequence 11892, A
C 40	15.2	76.0	89220	4	US-09-949-016-13655	Sequence 12655, A
C 41	15.2	76.0	89224	4	US-09-949-016-15572	Sequence 15572, A
C 42	15.2	76.0	98302	4	US-09-949-016-16847	Sequence 16847, A
C 43	15.2	76.0	99304	4	US-09-949-016-15440	Sequence 15440, A
C 44	15.2	76.0	103447	4	US-09-949-016-16320	Sequence 16320, A
C 45	15.2	76.0	119649	4	US-09-949-016-12537	Sequence 12537, A

ALIGNMENTS

RESULT 1

US-09-949-016-12174/c

; Sequence 12174, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12174

; LENGTH: 153642

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(153642)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12174

Query Match 84.0%; Score 16.8; DB 4; Length 153642;

Best Local Similarity 90.0%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAGGATT 20

DB 117870 GGAATGTTGAGTCAGGTTT 117851

RESULT 2

US-09-949-016-15635/c

; Sequence 15635, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15635
; LENGTH: 153643
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(153643)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15635

Query Match      84.0%; Score 16.8; DB 4; Length 153643;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGGATT 20
   ||||| ||||| ||||| ||||| |||||
Db 117870 GGAATGTTGAGTCAAGGTTT 117851

RESULT 3
US-09-621-976-17595
; Sequence 17595, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17595
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17595

Query Match      82.0%; Score 16.4; DB 4; Length 503;
Best Local Similarity 94.4%; Pred. No. 98;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGAT 19
   ||||| ||||| ||||| ||||| |||||
Db 68 GAAGATGAGTCAAGGAT 85

RESULT 4
US-09-641-638-552
; Sequence 552, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 552
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 439

; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 552
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: allele
; LOCATION: 439

; OTHER INFORMATION: 12-65-98 : polymorphic base C or T
; NAME/KEY: misc binding
; LOCATION: 419..438
; OTHER INFORMATION: 12-65-98.misl, potential
; NAME/KEY: misc binding
; LOCATION: 440..459
; OTHER INFORMATION: 12-65-98.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 342..360
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 777..797
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 427..451
; OTHER INFORMATION: 12-65-98 potential probe
; US-09-641-638-552

Query Match      82.0%; Score 16.4; DB 3; Length 939;
Best Local Similarity 94.4%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGGA 18
   ||||| ||||| ||||| ||||| |||||
Db 548 GGAAGTTGAGTCAAGCA 565

RESULT 5
US-10-170-097-552
; Sequence 552, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 552
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 439
```

OTHER INFORMATION: 12-65-98 : polymorphic base C or T  
FEATURE:  
NAME/KEY: misc\_binding  
LOCATION: 419..438  
OTHER INFORMATION: 12-65-98.mis1, potential  
FEATURE:  
NAME/KEY: misc\_binding  
LOCATION: 440..459  
OTHER INFORMATION: 12-65-98.mis2, potential complement  
FEATURE:  
NAME/KEY: primer\_bind  
LOCATION: 342..360  
OTHER INFORMATION: upstream amplification primer  
FEATURE:  
NAME/KEY: primer\_bind  
LOCATION: 777..797  
OTHER INFORMATION: downstream amplification primer, complement  
FEATURE:  
NAME/KEY: misc\_binding  
LOCATION: 427..451  
OTHER INFORMATION: 12-65-98 potential probe  
US-10-170-097-552

Query Match 82.0%; Score 16.4; DB 4; Length 939;  
Best Local Similarity 94.4%; Pred. No. 1e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGA 18  
Db 548 GGAAGGTTGAGTCAAGCA 565

## RESULT 6

US-09-949-016-12505  
Sequence 12505, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12505  
LENGTH: 670689  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(670689)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12505

Query Match 82.0%; Score 16.4; DB 4; Length 670689;  
Best Local Similarity 94.4%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAGGTTGAGTCAAGGAT 19  
Db 558453 GAAGGTTGATTCAGGAT 558470

## RESULT 7

US-09-949-016-14207  
Sequence 14207, Application US/09949016

Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14207  
LENGTH: 670690  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(670690)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14207

Query Match 82.0%; Score 16.4; DB 4; Length 670690;  
Best Local Similarity 94.4%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GAAGGTTGAGTCAAGGAT 19  
Db 558453 GAAGGTTGATTCAGGAT 558470

## RESULT 8

US-09-797-906-3/c  
Sequence 3, Application US/09797906  
Patent No. 6329188  
GENERAL INFORMATION:  
APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY  
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
USERS THEREOF  
FILE REFERENCE: CL001151CIP  
CURRENT APPLICATION NUMBER: US/09/797,906  
CURRENT FILING DATE: 2001-03-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 84495  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(84495)  
OTHER INFORMATION: n = A,T,C or G  
US-09-797-906-3

Query Match 80.0%; Score 16; DB 3; Length 84495;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAG 16  
Db 34427 GGAAGGTTGAGTCAAG 34412

## RESULT 9

US-09-949-016-16549  
Sequence 16549, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:

us-09-899-276c-8.rn1

Tue Aug 9 17:01:29 2005

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16549
; LENGTH: 47555
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16549

Query Match          79.0%; Score 15.8; DB 4; Length 47555;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GAAGTTGAGTCAAGGATT 20
      |||||
Db      33158 GAAGTTGAGTCAAGGATT 33176

RESULT 10
US-09-949-016-12598
; Sequence 12598, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12598
; LENGTH: 49673
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12598

Query Match          79.0%; Score 15.8; DB 4; Length 49673;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GAAGTTGAGTCAAGGATT 20
      |||||
Db      33158 GAAGTTGAGTCAAGGATT 33176

RESULT 11
US-09-949-016-15821
; Sequence 15821, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15821
; LENGTH: 50563
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15821

Query Match          79.0%; Score 15.8; DB 4; Length 50563;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GAAGTTGAGTCAAGGATT 20
      |||||
Db      21340 GAAAGTTGAGTTAAGGATT 21358

RESULT 12
US-09-741-150-3/c
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

Query Match          79.0%; Score 15.8; DB 3; Length 112132;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GAAGTTGAGTCAAGGATT 20
      |||||
Db      32191 GAAAGTTGAGTCAAGGATT 32173

RESULT 13
US-10-160-187-3/c
; Sequence 3, Application US/10160187
; Patent No. 6620607
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000968DIV
; CURRENT APPLICATION NUMBER: US/10/160,187
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/252,410
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 09/741,150
; PRIOR FILING DATE: 2000-12-21

```

; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 112132  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(112132)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-160-187-3

Query Match 79.0%; Score 15.8; DB 4; Length 112132;  
Best Local Similarity 89.5%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20  
||| ||||| ||||| |||||  
Db 32191 GAAAGTTGAGCAAGGATT 32173

## RESULT 14

US-09-949-016-15094  
; Sequence 15094, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15094

; LENGTH: 209210  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(209210)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-15094

Query Match 79.0%; Score 15.8; DB 4; Length 209210;  
Best Local Similarity 89.5%; Pred. No. 3.5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20  
||| ||||| ||||| |||||  
Db 25046 GAAGTTGAGTCAGGATT 25064

## RESULT 15

US-09-949-016-12465  
; Sequence 12465, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12465  
; LENGTH: 45299  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(45299)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12465

Query Match 77.0%; Score 15.4; DB 4; Length 45299;  
Best Local Similarity 94.1%; Pred. No. 4.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGG 17  
||||| ||||| ||||| |||||  
Db 36932 GGAAGTTGAGCCAAGG 36948

Search completed: August 4, 2005, 14:41:12  
Job time : 9.12146 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 08:04:16 ; Search time 35.6761 Seconds  
(without alignments)  
3633.986 Million cell updates/sec

Title: US-09-899-276C-8

Perfect score: 20

Sequence: 1 ggaaggttgagtcaggatt 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	20	100.0	20	9	US-09-899-276-8
2	17	85.0	21	9	Sequence 8, Appli
3	17	85.0	600	9	Sequence 10, Appl
4	17	85.0	11793	19	US-09-899-276-1
5	16.8	84.0	472	21	US-10-685-705-4
6	16.8	84.0	1817	19	Sequence 4, Appli
7	16.4	82.0	939	17	Sequence 816, App
					Sequence 16253, A
					Sequence 552, App

8	16.4	82.0	939	21	US-10-926-684-552	Sequence 552, App
9	15.8	79.0	201	19	US-10-741-601-9597	Sequence 9597, Ap
10	15.8	79.0	201	19	US-10-741-601-13271	Sequence 13271, A
11	15.8	79.0	201	21	US-10-741-600-23956	Sequence 23956, A
12	15.8	79.0	201	21	US-10-741-600-31551	Sequence 31551, A
13	15.8	79.0	366	18	US-10-424-599-130815	Sequence 130815, A
14	15.8	79.0	367	20	US-10-425-115-23154	Sequence 23154, A
15	15.8	79.0	394	20	US-10-425-115-84482	Sequence 84482, A
16	15.8	79.0	427	13	US-10-027-632-36716	Sequence 36716, A
17	15.8	79.0	427	17	US-10-027-632-36716	Sequence 36716, A
18	15.8	79.0	459	14	US-10-198-846-3203	Sequence 3203, Ap
19	15.8	79.0	477	18	US-10-424-599-34603	Sequence 34603, A
20	15.8	79.0	507	13	US-10-027-632-68331	Sequence 68331, A
21	15.8	79.0	507	13	US-10-027-632-302994	Sequence 302994, A
22	15.8	79.0	507	17	US-10-027-632-68331	Sequence 68331, A
23	15.8	79.0	507	13	US-10-027-632-302994	Sequence 302994, A
24	15.8	79.0	570	13	US-10-027-632-126479	Sequence 126479, A
25	15.8	79.0	570	17	US-10-027-632-126479	Sequence 126479, A
26	15.8	79.0	617	20	US-10-425-115-158788	Sequence 158788, A
27	15.8	79.0	648	18	US-10-424-599-134677	Sequence 134677, A
28	15.8	79.0	758	14	US-10-198-846-11678	Sequence 11678, A
29	15.8	79.0	820	18	US-10-425-114-29383	Sequence 29383, A
30	15.8	79.0	1281	18	US-10-424-599-53452	Sequence 53452, A
31	15.8	79.0	1385	18	US-10-425-114-7982	Sequence 7982, Ap
32	15.8	79.0	2133	20	US-10-739-930-3532	Sequence 3532, Ap
33	15.8	79.0	2595	18	US-10-424-599-130251	Sequence 130251, A
34	15.8	79.0	4251	19	US-10-437-963-56906	Sequence 56906, A
35	15.8	79.0	6675	21	US-10-481-582-7	Sequence 7, Appli
36	15.8	79.0	96587	11	US-09-397-722-250	Sequence 250, App
37	15.8	79.0	97415	19	US-10-322-281-287	Sequence 287, App
38	15.8	79.0	112132	9	US-09-741-150-3	Sequence 3, Appli
39	15.8	79.0	112132	20	US-10-612-012-3	Sequence 3, Appli
40	15.8	79.0	112486	19	US-10-741-601-5641	Sequence 5641, Ap
41	15.8	79.0	112486	21	US-10-741-600-17642	Sequence 17642, A
42	15.8	79.0	161700	19	US-10-741-601-5623	Sequence 5623, Ap
43	15.8	79.0	161700	21	US-10-741-600-17590	Sequence 17590, A
44	15.8	79.0	312477	19	US-10-317-883A-12	Sequence 12, Appl
45	15.8	79.0	402850	10	US-09-844-653-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-899-276-8  
; Sequence 8, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Delius, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899, 276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Part of 3'-DHRS  
US-09-899-276-8

Query Match 100.0%; Score 20; DB 9; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGGATT 20  
|||||  
Db 1 GGAAGTTGAGTCAAGGATT 20

## RESULT 2

US-09-899-276-10  
; Sequence 10, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Delius, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899,276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Oligonucleotide  
US-09-899-276-10

Query Match 85.0%; Score 17; DB 9; Length 21;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGG 17  
|||||  
Db 1 GGAAGTTGAGTCAAGG 17

## RESULT 3

US-09-899-276-1  
; Sequence 1, Application US/09899276  
; Patent No. US20020106355A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosl, Frank  
; APPLICANT: Soto, Ubaldo  
; APPLICANT: Coy, Johannes  
; APPLICANT: Finzer, Patrick  
; APPLICANT: Delius, Hajo  
; APPLICANT: Poustka, Annemarie  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Patzelt, Andrea  
; TITLE OF INVENTION: No. US20020106355A1el Regulatory Sequences of the MCP-1 Gene  
; FILE REFERENCE: 012627-023  
; CURRENT APPLICATION NUMBER: US/09/899,276  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: EP 00 114 560.6  
; PRIOR FILING DATE: 2000-07-06  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-899-276-1

Query Match 85.0%; Score 17; DB 9; Length 600;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGG 17  
|||||  
Db 158 GGAAGTTGAGTCAAGG 174

## RESULT 4

US-10-685-705-4  
; Sequence 4, Application US/10685705  
; Publication No. US20040177387A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Kentucky Research Foundation  
; APPLICANT: JAYAKRISHNA, Ambati  
; TITLE OF INVENTION: Methods And Animal Model For Analyzing Age-Related Macular  
; TITLE OF INVENTION: Degeneration  
; FILE REFERENCE: 050229-0415  
; CURRENT APPLICATION NUMBER: US/10/685,705  
; CURRENT FILING DATE: 2003-10-16  
; PRIOR APPLICATION NUMBER: 60/422,096  
; PRIOR FILING DATE: 2002-10-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 4  
; LENGTH: 11793  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-685-705-4

Query Match 85.0%; Score 17; DB 19; Length 11793;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGG 17  
|||||  
Db 10136 GGAAGTTGAGTCAAGG 10152

## RESULT 5

US-10-950-009-816/c  
; Sequence 816, Application US/10950009  
; Publication No. US20050069934A1  
; GENERAL INFORMATION:  
; APPLICANT: BERKA, Randy  
; APPLICANT: BASHKIROVA, Elena  
; APPLICANT: REY, Michael  
; TITLE OF INVENTION: Methods For Monitoring Gene Expression  
; FILE REFERENCE: 10541.200-US  
; CURRENT APPLICATION NUMBER: US/10/950,009  
; CURRENT FILING DATE: 2004-09-24  
; PRIOR APPLICATION NUMBER: 60/506,140  
; PRIOR FILING DATE: 2003-09-25  
; NUMBER OF SEQ ID NOS: 1190  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 816  
; LENGTH: 472  
; TYPE: DNA  
; ORGANISM: Trichoderma reesei  
US-10-950-009-816

Query Match 84.0%; Score 16.8; DB 21; Length 472;  
Best Local Similarity 90.0%; Pred. No. 1.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGGATT 20  
|||||  
Db 351 GGAAGTTGAGTCAAGGTTT 332

## RESULT 6

US-10-437-963-16253

Qy 1 GGAAGGTTGAGTCAAGGA 18  
|||  
Db 548 GGAAGGTTGAGTCAAGGA 565

RESULT 8  
US-10-926-684-552  
; Sequence 552, Application US/10926684  
; Publication No. US20050014190A1

; Sequence 532, Application OS/10922  
 ; Publication No. US20050014190A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Blumenfeld, Marta

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

; LOCATION: 342...360  
; OTHER INFORMATION: upstream amplification primer  
; FEATURE:  
; NAME/KEY: primer bind  
; LOCATION: 777..797  
; OTHER INFORMATION: downstream amplification primer, complement  
; FEATURE:  
; NAME/KEY: misc binding  
; LOCATION: 427..451  
; OTHER INFORMATION: 12-65-98 potential probe  
US-10-926-684-552

Query Match 82.0%; Score 16.4; DB 21; Length 939;  
Best Local Similarity 94.4%; Pred. No. 2.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGAAGTTGAGTCAAGGA 18  
|||||  
DB 548 GGAAGTTGAGTCAAGCA 565

## RESULT 9

US-10-741-601-9597  
; Sequence 9597, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9597  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-9597

Query Match 79.0%; Score 15.8; DB 19; Length 201;  
Best Local Similarity 89.5%; Pred. No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20  
|||||  
DB 138 GAAGGTGAGTCAATGATT 156

## RESULT 10

US-10-741-601-13271/c  
; Sequence 13271, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13271  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-13271

Query Match 79.0%; Score 15.8; DB 19; Length 201;  
Best Local Similarity 89.5%; Pred. No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20

Db 64 GAAGATGAGTCATGGATT 46  
|||||

## RESULT 11

US-10-741-600-23956  
; Sequence 23956, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 23956  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-23956

Query Match 79.0%; Score 15.8; DB 21; Length 201;  
Best Local Similarity 89.5%; Pred. No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20  
|||||  
DB 138 GAAGATGAGTCATGGATT 156

## RESULT 12

US-10-741-600-31551/c  
; Sequence 31551, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 31551  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-31551

Query Match 79.0%; Score 15.8; DB 21; Length 201;  
Best Local Similarity 89.5%; Pred. No. 4.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GAAGTTGAGTCAAGGATT 20  
|||||  
DB 64 GAAGATGAGTCATGGATT 46

## RESULT 13

US-10-424-599-130815/c  
; Sequence 130815, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 130815  
; LENGTH: 366  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_89133C.1  
US-10-424-599-130815

Query Match 79.0%; Score 15.8; DB 18; Length 366;  
Best Local Similarity 89.5%; Pred. No. 5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGAT 19  
|||||  
DB 108 GGAAGGTAGATTAAAGGAT 90

## RESULT 14

US-10-425-115-23154/c  
; Sequence 23154, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 23154  
; LENGTH: 367  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_121122C.1  
US-10-425-115-23154

Query Match 79.0%; Score 15.8; DB 20; Length 367;  
Best Local Similarity 89.5%; Pred. No. 5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGAT 19  
|||||  
DB 294 GGAAGGTTGCCCAAGGAT 276

## RESULT 15

US-10-425-115-84482/c  
; Sequence 84482, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 84482  
; LENGTH: 394  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_177058C.1  
US-10-425-115-84482

Query Match 79.0%; Score 15.8; DB 20; Length 394;  
Best Local Similarity 89.5%; Pred. No. 5e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGAAGGTTGAGTCAAGGAT 19  
|||||  
DB 380 GGAAGGTAGAGTCAAGGAT 362

Search completed: August 4, 2005, 15:54:30  
Job time : 37.6761 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 16:14:42 ; Search time 996 Seconds  
(without alignments)  
7881.270 Million cell updates/sec

Title: SEQ13-8698-9073-50N-TGATCA

Perfect score: 162

Sequence: 1 cttcaagaccattgtggcca.....nnnnnnnnnnntgatca 162

Scoring table: IDENTITY\_NUCDX

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	157	96.9	360	6	CQ069701
C 2	157	96.9	360	6	CQ096880
C 3	157	96.9	360	6	CQ135681
C 4	157	96.9	360	6	CQ174058
C 5	157	96.9	360	6	CQ218977
C 6	157	96.9	360	6	CQ257544
C 7	157	96.9	360	6	CQ294784
C 8	157	96.9	360	6	CQ331440
C 9	157	96.9	478	6	AX563646
10	157	96.9	478	9	HUMSECP3
11	157	96.9	482	6	AX774781
12	157	96.9	482	9	HUMMSPA3
13	157	96.9	647	6	AR567996
14	157	96.9	647	6	AX698741
15	157	96.9	661	6	AR559337
16	157	96.9	661	6	BD080551
17	157	96.9	661	9	S69738
18	157	96.9	718	9	HUMSECPA
19	157	96.9	725	6	AR337874

20	157	96.9	725	6	AR380799
21	157	96.9	725	6	AX774742
22	157	96.9	725	9	HS MCP1
23	157	96.9	729	6	CQ724136
24	157	96.9	739	6	AX577974
25	157	96.9	739	9	S71513
26	157	96.9	741	6	A17786
27	157	96.9	741	6	AR094465
28	157	96.9	741	6	E05611
29	157	96.9	741	6	AR380620
30	157	96.9	741	9	HUMMCAF
31	157	96.9	743	9	BC009716
32	157	96.9	757	6	CQ831706
33	157	96.9	1712	6	BD195644
34	157	96.9	1712	6	CQ775689
35	157	96.9	1712	6	AR352699
36	157	96.9	1822	6	BD195643
37	157	96.9	1822	6	CQ775688
38	157	96.9	1822	6	AR352698
39	157	96.9	2243	9	HSJEPR
40	157	96.9	2776	9	HUMMCHMP
41	157	96.9	3174	9	AF519531
42	157	96.9	11793	9	HSY18933
43	157	96.9	147416	9	AC005549
44	156	96.3	554	6	E05599
45	153.8	94.9	786	11	BV166059

## ALIGNMENTS

RESULT 1  
CQ069701/C  
LOCUS CQ069701 360 bp DNA linear PAT 20-JAN-2004  
DEFINITION Sequence 5501 from Patent WO0157278.  
ACCESSION CQ069701  
VERSION CQ069701.1 GI:41039570

KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human hela cells or other human cervical epithelial cells  
Patent: WO 0157278-A 5501 09-AUG-2001;  
Aecomica, Inc. (US)

## FEATURES

Location/Qualifiers  
source 1..360  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AC005549.1-EXPRESSED IN HELA, SIGNAL = 2.1"

## ORIGIN

Query Match 96.9%; Score 157; DB 6; Length 360;  
Best Local Similarity 68.2%; Pred. No. 1.7e-24;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTGTGGCCAGGAGATCTGTGCTGACCCCAAGCAAGAGTGGTTTCAGGA 60  
Db 358 CTTCAAGACCATTGTGGCCAGGAGATCTGTGCTGACCCCAAGCAAGAGTGGTTTCAGGA 299  
QY 61 TTCCATGACCACCTGGACCAAGCAAAACCCAAAGCTTCGANNNNNNNNNNNN 120  
Db 298 TTCCATGACCACCTGGACCAAGCAAAACCCAAAGCTTCGACACTCCTCCACA 239  
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157  
Db 238 ACCCAAGAATCTGACGCTAACTTATTTCCCTAGCT 202

```

RESULT 2
LOCUS      CQ096880/c          360 bp      DNA          linear          PAT 21-JAN-2004
DEFINITION Sequence 5739 from Patent WO0157272.
ACCESSION  CQ096880
VERSION     CQ096880.1  GI:41065906
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE       Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human placenta
JOURNAL     Patent: WO 0157272-A 5739 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES
            Location/Qualifiers
            1..360
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="MAP TO AC005549.1-EXPRESSED IN PLACENTA, SIGNAL =
            2.9"

ORIGIN
Query Match      96.9%; Score 157; DB 6; Length 360;
Best Local Similarity 68.2%; Pred. No. 1.7e-24;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY      1  CTTCAAGACCAATTGTGGCCAAAGAGATCTGTCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60
Db      358 CTTCAAGACCAATTGTGGCCAAAGAGATCTGTCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 299

QY      61  TTCCATGGACCACTCTGGACAAGCAACCAACCTCGAAGACTTGGANNNNNNNNNNNN 120
Db      298 TTCCATGGACCACTCTGGACAAGCAACCAACCTCGAAGACTTGGAACTCACTCCACA 239

QY      121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db      238 ACCCAAGAATCTGCAGTAACTTATTTCCCTAGCT 202

RESULT 4
LOCUS      CQ174058/c          360 bp      DNA          linear          PAT 21-JAN-2004
DEFINITION Sequence 5454 from Patent WO0157274.
ACCESSION  CQ174058
VERSION     CQ174058.1  GI:41168794
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE       Human genome-derived single exon nucleic acid probes useful for
            analysis of gene expression in human heart
JOURNAL     Patent: WO 0157274-A 5454 09-AUG-2001;
            Aeomica, Inc. (US)
FEATURES
            Location/Qualifiers
            1..360
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="MAP TO AC005549.1-EXPRESSED IN HEART, SIGNAL = 2.6"

ORIGIN
Query Match      96.9%; Score 157; DB 6; Length 360;
Best Local Similarity 68.2%; Pred. No. 1.7e-24;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY      1  CTTCAAGACCAATTGTGGCCAAAGAGATCTGTGTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60
Db      358 CTTCAAGACCAATTGTGGCCAAAGAGATCTGTGTGACCCCAAGCAGAGAGTGGGTTTCAGGA 299

QY      61  TTCCATGGACCACTCTGGACAAGCAACCAACCTCGAAGACTTGGANNNNNNNNNNNN 120
Db      298 TTCCATGGACCACTCTGGACAAGCAACCAACCTCGAAGACTTGGAACTCACTCCACA 239

QY      121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db      238 ACCCAAGAATCTGCAGTAACTTATTTCCCTAGCT 202

RESULT 5
LOCUS      CQ218977/c          360 bp      DNA          linear          PAT 21-JAN-2004
DEFINITION Sequence 5816 from Patent WO0157273.
ACCESSION  CQ218977
VERSION     CQ218977.1  GI:41200544
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE       HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
            ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
            3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
            60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03

ORIGIN
Query Match      96.9%; Score 157; DB 6; Length 360;
Best Local Similarity 68.2%; Pred. No. 1.7e-24;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

```

Db 238 ACCCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 202  
 RESULT 7  
 LOCUS CQ294784/c  
 DEFINITION CQ294784  
 Sequence 5889 from Patent WO0186003.  
 ACCESSION CQ294784  
 VERSION CQ294784.1 GI:41255361  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
 Human genome-derived single exon nucleic acid probes useful for  
 analysis of gene expression in human lung  
 Patent: WO 0186003-A 5889 15-NOV-2001;  
 aeomica, Inc. (US)  
 FEATURES  
 Location/Qualifiers  
 1..360  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="MAP TO AC005549.1-EXPRESSED IN LUNG, SIGNAL = 2.5"  
 ORIGIN  
 Query Match 96.9%; Score 157; DB 6; Length 360;  
 Best Local Similarity 68.2%; Pred. No. 1.7e-24;  
 Matches 107; Conservative 50; Mismatch 0; Indels 0; Gaps 0;

QY	1	CTTCAAGACCATTTGTGGCCCAAGAGATCTGTGCTACCCCAAGCAGAAGTGGGTTTCAGGA	60
Db	358	CTTCAAGACCATTTGTGGCCCAAGAGATCTGTGCTACCCCAAGCAGAAGTGGGTTTCAGGA	299
QY	61	TTTCCATGGACCACTCGGCAAGCAACCAAACTCGAGACTTCGANNNNNNNNNNNN	120
Db	298	TTTCCATGGACCACTCGGCAAGCAACCAAACTCGAGACTTCGAACTTGAACACTCACTCCACA	239
QY	121	NN	157
Db	238	ACCCAAGACTTCGACCTAACTATTATTTCCCTAGCT	202

  

RESULT	8
LOCUS	QO3311440/c
DEFINITION	Sequence 5534 from Patent WO0157275.
ACCESSION	QO3311440
VERSION	QO3311440.1
KEYWORDS	GI:41280007
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE	Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human brain
JOURNAL	Patent: WO 0157275-A 5534 09-AUG-2001; Aeomica, Inc. (US)
FEATURES	Location/Qualifiers
source	1..360 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="MAP TO AC005549.1-EXPRESSED IN BRAIN, SIGNAL = 2.5"

  

Query Match	96.9%;	Score 157;	DB 6;	Length 360;
Best Local Similarity	68.2%;	Pred. No. 1.7e-24;		
Matches 107;	Conservative 50;	Mismatches 0;	Indels 0;	Gaps 0;

  

ORIGIN	
--------	--

Tue Aug 9 17:01:22 2005

```
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
Db 358 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 299
QY 61 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGGANNNNNNNNNNNN 120
Db 298 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGGAACTCACTCCACA 239
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 238 ACCCAAGAATCTGCAGCTAACTTATTTCCCTAGCT 202

RESULT 9
AX663646
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Oellers,N., Gehrmann,M., Kallabis,H., Hall,R., Schulze,T. and
Kroegel,C.
TITLE
Genes and proteins for prevention, prediction, diagnosis, prognosis
and treatment of chronic lung disease
JOURNAL
Patent: WO 02097127-A 21 05-DEC-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..478
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 96.9%; Score 157; DB 6; Length 478;
Best Local Similarity 68.2%; Pred. No. 1.8e-24;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
Db 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
QY 61 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGGANNNNNNNNNNNN 120
Db 61 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGAACCCACTCCACA 120
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 121 ACCCAAGAATCTGCAGCTAACTTATTTCCCTAGCT 157

RESULT 10
HUMSECP3
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Rollins,B.J., Stier,P., Ernst,T. and Wong,G.G.
TITLE
The human homolog of the JE gene encodes a monocyte secretory
protein
JOURNAL
Mol. Cell. Biol. 9 (11), 4687-4695 (1989)

MEDLINE
90097880
PUBMED
2513477
REFERENCE
2 (bases 1 to 478)
AUTHORS
Rollins,B.J., Stier,P., Ernst,T. and Wong,G.G.
JOURNAL
Unpublished (1989)
COMMENT
Original source text: Human lung fibroblast WI38 cell line, DNA,
clones hJE-34 and lambda-hJE-7.
[2] sites for [1].
Draft entry and computer readable sequence for [1] kindly provided
by B.J.Rollins 17-JUL-1989.
FEATURES
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/cell_line="WI38"
/cell_type="fibroblast"
/tissue_type="lung"
join(M28223.1:532..675,M28224.1:1..118,1..478)
/genes="JE"
join(M28223.1:600..675,M28224.1:1..118,1..106)
/genes="JE"
/codon_start=1
/product="monocyte secretory protein"
/protein_id="AAA60308.1"
/db_xref="GI:338007"
/translations="MKVSAALCLLLIAATFIQGLAQPDAINAPVTCYVFNTRKIS
VQLASYRRITSSKCPKEAVIFKTIVAKEICADPKQKWQVDSMDHLDKQTQTKT"
mat_peptide
join(M28223.1:669..675,M28224.1:1..118,1..103)
/genes="JE"
/product="monocyte secretory protein"
exon
1..478
/genes="JE"
/number=3

ORIGIN
Query Match 96.9%; Score 157; DB 9; Length 478;
Best Local Similarity 68.2%; Pred. No. 1.8e-24;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
Db 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
QY 61 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGGANNNNNNNNNNNN 120
Db 61 TTCCATGACCACTTGCGCAAGCAAAACCCAACTCCGAAGACTTGAACCCACTCCACA 120
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 121 ACCCAAGAATCTGCAGCTAACTTATTTCCCTAGCT 157

RESULT 11
AX774781
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Raponi,M.
TITLE
Methods for assessing and treating leukemia
JOURNAL
Patent: WO 0303129-A 97 08-MAY-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
Location/Qualifiers
1..482
/organism="Homo sapiens"
/mol_type="unassigned DNA"

RESULT 12
AX774781
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Raponi,M.
TITLE
Methods for assessing and treating leukemia
JOURNAL
Patent: WO 0303129-A 97 08-MAY-2003;
Ortho-Clinical Diagnostics, Inc. (US)
FEATURES
Location/Qualifiers
1..482
/organism="Homo sapiens"
/mol_type="unassigned DNA"
```

Query Match	96.9%;	Score 157;	DB 9;	Length 482;
Best Local Similarity	68.2%;	Pred. No. 1.8e-24;		
Matches 107;	Conservative	50;	Mismatches 0;	Indels 0;
Gaps	0;			

  

Qy	1	CTTCAAGACCAATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAACTGGGGTTTCAGGA	60
Db	6	CTTCAAGACCAATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAACTGGGGTTTCAGGA	65
Qy	61	TTCCATGGGACCACTGTGACCAAGCAAAACCCAAACTCCGAAGACTTGANNNNNNNNNNNN	120
Db	66	TTCCATGGGACCACTGTGACCAAGCAAAACCCAAACTCCGAAGACTTGAACTTCATCCACCA	125

Query Match 96.9%; Score 157; DB 6; Length 647;  
Best Local Similarity 68.2%; Pred. No. 1.9e-24;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATGTGGCCAAAGGACATCTGTCTCACCCCAAGCAGAAGTGGGTTCAAGGA 60

Db 195 CTTCAAGACCATTTGTGGCAAGGAGATCTGTGCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 254  
Qy 61 TTCCATGGACCACTGGACAGCAAAACCCAAACTCCGAAGACTTGANNNNNNNNNNNN 120  
Db 255 TTCCATGGACCACTGGACAGCAAAACCCAAACTCCGAAGACTTGAAACACTCACTCCACA 314  
Qy 121 NNN 157  
Db 315 ACCCAAGAATCTGCAGCTAACTTATTTCCTCCCTAGCT 351

RESULT 15  
AR559337  
LOCUS AR559337 661 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 104 from patent US 6750015.  
ACCESSION AR559337  
VERSION AR559337.1 GI:53968753  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 661)  
AUTHORS Horwitz,K.B. and Richer,J.  
TITLE Progesterone receptor-regulated gene expression and methods related thereto  
JOURNAL Patent: US 6750015-A 104 15-JUN-2004;  
FEATURES  
source Location/Qualifiers  
1..661  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 96.9%; Score 157; DB 6; Length 661;  
Best Local Similarity 68.2%; Pred. No. 1.9e-24;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTTCAAGACCATTTGTGGCCAAAGAGATCTGTGTCGACCCCAAGCAGAAAGTGGGTTTCAGGA 60  
Db 226 CTTCAAGACCATTTGTGGCCAAAGAGATCTGTGTCGACCCCAAGCAGAAAGTGGGTTTCAGGA 285  
Qy 61 TTCCATGGACCACTGGACAGCAAAACCCAAACTCCGAAGACTTGANNNNNNNNNNNNNNNN 120  
Db 286 TTCCATGGACCACTGGACAGCAAAACCCAAACTCCGAAGACTTGAAACACTCACTCCACA 345  
Qy 121 NNN 157  
Db 346 ACCCAAGAATCTGCAGCTAACTTATTTCCTCCCTAGCT 382

Search completed: August 4, 2005, 17:10:56  
Job time : 1003 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 15:58:12 ; Search time 220 seconds  
(without alignments)  
4359.082 Million cell updates/sec

Title: SEQ13-8698-9073-50N-TGATCA

Perfect score: 162

Sequence: 1 cttcaagaccattgtgccca.....nnnnnnnnnnnnntgatca 162

Scoring table: IDENTITY NUCDX

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001s:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002bs:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	157	96.9	360	AAI15568	AAI15568 Probe #55
C 2	157	96.9	360	ABA57500	ABA57500 Human foe
C 3	157	96.9	360	AAI37053	AAI37053 Probe #57
C 4	157	96.9	360	ABA26988	ABA26988 Probe #54
C 5	157	96.9	360	AAK31146	AAK31146 Human bon
C 6	157	96.9	360	AAK05543	AAK05543 Human bra
C 7	157	96.9	360	ABS30826	ABS30826 Human liv
C 8	157	96.9	360	ABS05898	ABS05898 Human gen
C 9	157	96.9	473	ABV96715	ABV96715 Human pan
C 10	157	96.9	475	ABV97697	ABV97697 Human pan
C 11	157	96.9	478	ACC46770	ACC46770 Human COP
C 12	157	96.9	478	ADN95619	ADN95619 Human BEC
C 13	157	96.9	478	ADP13526	ADP13526 Renal cel
C 14	157	96.9	482	ADE84878	ADE84878 Farnesyl
C 15	157	96.9	647	ABS68800	ABS68800 Human mon
C 16	157	96.9	661	AAA74882	AAA74882 Human che
C 17	157	96.9	725	AAQ85370	AAQ85370 Chemoattr
C 18	157	96.9	725	AAAX80631	AAAX80631 Monocyte
C 19	157	96.9	725	AAAD58817	AAAD58817 Human mon
C 20	157	96.9	725	ADD18494	ADD18494 Human pro

21	157	96.9	725	10	ADE84839	Ad84839 Farnesyl
22	157	96.9	725	11	ADI32018	ADI32018 Human CDN
23	157	96.9	725	13	ADR05375	Adr05375 Human mon
24	157	96.9	738	1	AAN91337	Aan91337 DNA which
25	157	96.9	739	2	AAV10341	Aav10341 cDNA enco
26	157	96.9	739	6	ABV78071	Abv78071 Hypoxia-r
27	157	96.9	739	8	ABZ34738	Abz34738 Coding se
28	157	96.9	739	12	ADF45446	Adf45446 Human vas
29	157	96.9	739	12	ADN04745	Adn04745 Antipsori
30	157	96.9	739	13	ADR24988	Adr24988 Breast ca
31	157	96.9	739	13	ADP24778	Adp24778 PRO poly
32	157	96.9	741	2	AAQ30748	Aaq30748 pMCF7. 3
33	157	96.9	741	3	AAA34899	Aaa34899 Human ade
34	157	96.9	741	3	AAF21021	Aaf21021 Human low
35	157	96.9	741	10	ABZ96715	Abz96715 Human nuc
36	157	96.9	741	11	ADI31839	Adi31839 Human CDN
37	157	96.9	741	11	ABD20564	Abd20564 Human pul
38	157	96.9	743	12	ADO03801	Ado03801 Human Cc1
39	157	96.9	756	10	ADI02669	Adi02669 Human CDN
40	157	96.9	757	10	ADD14996	Add14996 Human mon
41	157	96.9	757	11	ADN95635	Adn95635 Human BEC
42	157	96.9	757	12	ADO03800	Ado03800 Human Cc1
43	157	96.9	757	12	ADP75913	Adp75913 Human SCF
44	157	96.9	757	12	ADQ76211	Adq76211 Chemokine
45	157	96.9	757	12	ADQ09331	Adq09331 Human SCF

## ALIGNMENTS

RESULT 1

AAI15568/c

ID AAI15568 standard; DNA; 360 BP.

XX AAI15568;

XX

DT 12-OCT-2001 (first entry)

XX

DE Probe #5501 for gene expression analysis in human cervical cell sample.

XX

KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX

OS Homo sapiens.

XX

PN WO200157278-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000670.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR

PR 26-MAY-2000; 2000US-0207458P.

PR

PR 30-JUN-2000; 2000US-00608408.

PR

PR 03-AUG-2000; 2000US-00632366.

PR

PR 21-SEP-2000; 2000US-0234687P.

PR

PR 27-SEP-2000; 2000US-0236359P.

PR

PR 04-OCT-2000; 2000GB-00024263.

XX

(MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488901/53.

XX

Human genome-derived single exon nucleic acid probes useful for analyzing

gene expression in human cervical epithelial cells.

XX

PS Claim 25; SEQ ID NO 5501; 487pp; English.

XX

CC The present invention relates to human single exon nucleic acid probes

(SENPs). The present sequence is one such probe. The SENPs are derived

from human HeLa cells. The SENPs can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging of  
CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match 96.9%; Score 157; DB 4; Length 360;  
Best Local Similarity 68.2%; Pred. No. 0.73;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGAGTGGTTTCAGGA 60  
DB 358 CTTCAAGACCATTTGTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGAGTGGTTTCAGGA 299

QY 61 TTCCATGGACCACTGGACAGCAACCCAAACTCCGAAGACTTGANNNNNNNNNNNN 120  
DB 298 TTCCATGGACCACTGGACAGCAACCCAAACTCCGAAGACTTGAAACACTCACTCCACA 239

QY 121 NNN 157

DB 238 ACCCAAGAATCTGCAGTAACCTATTATTTCCCTAGCT 202

## RESULT 2

ABAS7500/c  
ID ABAS7500 standard; DNA; 360 BP.

XX AC ABAS7500;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #5805.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human fetal liver.

XX PS Claim 1; SEQ ID NO 5805; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring  
XX CC human gene expression in a sample derived from human foetal liver. The  
XX CC single exon nucleic acid probes may be used for predicting, measuring and  
XX CC displaying gene expression in samples derived from human fetal liver. The  
XX CC present sequence is a single exon nucleic acid probe of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match 96.9%; Score 157; DB 4; Length 360;

Best Local Similarity 68.2%; Pred. No. 0.73;

Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
DB 358 CTTCAAGACCATTTGTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 299

QY 61 TTCCATGGACCACTGGACAGCAACCCAAACTCCGAAGACTTGANNNNNNNNNNNN 120  
DB 298 TTCCATGGACCACTGGACAGCAACCCAAACTCCGAAGACTTGAAACACTCACTCCACA 239

QY 121 NNN 157

DB 238 ACCCAAGAATCTGCAGTAACCTATTATTTCCCTAGCT 202

## RESULT 3

AAI37053/c

ID AAI37053 standard; DNA; 360 BP.

XX AC AAI37053;

XX DT 17-OCT-2001 (first entry)

XX DE Probe #5739 used to measure gene expression in human placenta sample.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW Genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488997/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human placenta.

XX PS Claim 25; SEQ ID NO 5739; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENPs).  
XX CC The present sequence is one such probe. The probes are useful for  
XX CC producing a microarray for predicting, measuring and displaying gene  
XX CC expression in samples derived from human placenta. The probes are useful  
XX CC for antenatal diagnosis of human genetic disorders

XX SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;

Query Match 96.9%; Score 157; DB 4; Length 360;

Best Local Similarity 68.2%; Pred. No. 0.73;

Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
Db 358 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 299
QY 61 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 298 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGAAACACTCACTCCACA 239
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db 238 ACCCAAGAATCTGCAGTAACCTTATTTTCCCTAGCT 202
```

```
RESULT 4
ABA26988/c
ID ABA26988 standard; DNA; 360 BP.
XX
AC ABA26988;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #5454 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488999/53.
XX
Single exon nucleic acid probes for analyzing gene expression in human hearts.
```

```
PS Claim 1; SEQ ID NO 5454; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;
```

```
Query Match 96.9%; Score 157; DB 4; Length 360;
Best Local Similarity 68.2%; Pred. No. 0.73;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
```

```
Db 358 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 299
QY 61 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 298 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGAAACACTCACTCCACA 239
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db 238 ACCCAAGAATCTGCAGTAACCTTATTTTCCCTAGCT 202
```

```
RESULT 5
AAK31146/c
ID AAK31146 standard; DNA; 360 BP.
XX
AC AAK31146;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 5703.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488900/53.
XX
Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow.
```

```
PS Example 4; SEQ ID NO 5703; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;
```

```
Query Match 96.9%; Score 157; DB 4; Length 360;
Best Local Similarity 68.2%; Pred. No. 0.73;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
Db 358 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 299
QY 61 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 298 TTCCATGGACCACTGGACCAAGCAAAACCCAACTCCGAAGACTTGAAACACTCACTCCACA 239
```

RESULT 7
ABS30826/c
ID ABS30826 standard; DNA; 360 BP.
XX AC
XX ABS30826;
XX
XX 25-FEB-2003 (first entry)
XX Human liver single exon probe, SEQ ID No 5816.
DE
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX OS Homo sapiens.
XX WO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US000664.
PF
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488998/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
PT
XX Claim 1; SEQ ID NO 5816; 658pp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;
SQ
Query Match 96.9%; Score 157; DB 4; Length 360;
Best Local Similarity 68.2%; Pred. No. 0.73;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0
QY 1 CTTCAAGACCATGTGGCCAAAGGAGATCTGTCGACCCTTGAAGCAGCAAGTGCGTTCAGGA 60
DB 358 CTTCAAGACCATGTGGCCAAAGGAGATCTGTCGACCCTTGAAGCAGCAAGTGCGTTCAGGA 299
QY 61 TTCCATTGGACCACTTGGCAAGCAACACCCAAACTCCGAAGACCTTGCANNNNNNNNNNN 120
DB 298 TTCCATTGGACCACTTGGCAAGCAACACCCAAACTCCGAAGACCTTGCANNNNNNNNNNN 157
QY 121 NNN 120
DB 238 ACCCAAGAATCGACGTAACCTATTTTCCCCTAGCT 202

QY	121	NNNT	157
		.....	
Db	238	ACCCAGAATCTGCAGCTAACTTATTTCCTTAGCT	202
 RESULT 6			
AAK05543/c			
ID	AAK05543	standard; DNA; 360 BP.	
XX	AAK05543;		
XX			
DT	05-NOV-2001	(first entry)	
XX			
DE	Human brain expressed single exon probe SEQ ID NO: 5534.		
DE	Human; brain expressed exon; gene expression analysis; probe; microarray;		
KW	Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;		
KW	ss.		
OS	Homo sapiens.		
XX			
PN	WO200157275-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	30-JAN-2001; 2001WO-US000667.		
XX			
PR	04-FEB-2000; 2000US-0180312P.		
PR	26-MAY-2000; 2000US-0207456P.		
PR	30-JUN-2000; 2000US-00608408.		
PR	03-AUG-2000; 2000US-00632366.		
PR	21-SEP-2000; 2000US-0234687P.		
PR	27-SEP-2000; 2000US-0236359P.		
PR	04-OCT-2000; 2000GB-00024263.		
XX			
PA	(MOLE-) MOLECULAR DYNAMICS INC.		
XX			
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX			
DR	WPI; 2001-483446/52.		
XX			
PT	Single exon nucleic acid probes for analyzing gene expression in human		
PT	brains.		
XX			
PS	Example 4; SEQ ID NO 5534; 650pp + Sequence Listing; English.		
XX			
CC	The present invention provides a number of single exon nucleic acid		
CC	probes which are derived from genomic sequences expressed in the human		
CC	brain. They can be used to measure gene expression in brain cell samples,		
CC	which may enable the diagnosis and improved treatment of nervous system		
CC	diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,		
CC	epilepsy and cancers. The present sequence is one of the probes of the		
CC	invention		
XX			
SQ	Sequence 360 BP; 121 A; 65 C; 77 G; 97 T; 0 U; 0 Other;		
	Query Match	96.9%; Score 157; DB 4; Length 360;	
	Best Local Similarity	68.2%; Pred. No. 0.73;	
	Matches 107; Conservative	50; Mismatches 0; Indels 0; Gaps 0	
QY	1	CTTCAAGACCATTTGTGGCCAAGAGATCTGTCGACCCTCAAGCAGAGTGCGGTTCAGGA	60
Db	358	CTTCAAGACCATTTGTGGCCAAGAGATCTGTCGACCCTCAAGCAGAGTGCGGTTCAGGA	299
QY	61	TTCATGGACCACTCGACAAAGCAAACCCAAAACCTCGAAGACTTGANNNNNNNNNNNNN	120
Db	298	TTCATGGACCACTCGACAAAGCAAACCCAAAACCTCGAAGACTTGACACTCACTCCACA	239
QY	121	NNNT	157
		.....	
Db	238	ACCCAAGAATCTGCAGCTAACTTATTTCCTTAGCT	202



PT cancer.  
XX Claim 1; SEQ ID NO 2123; 300pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)  
CC complements of (a); (c) sequences consisting of at least 20 contiguous  
CC residues of (a); (d) sequences that hybridize to (a), under moderately  
CC stringent conditions; (e) sequences having at least 75% or 90% identity  
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
CC ABP68637) encoded by (I) and oligonucleotides can be used to detect cancer  
CC in a patient and compositions comprising polypeptides, polynucleotides,  
CC antibodies, fusion proteins, T cell populations and antigen presenting  
CC cells expressing the polypeptide are useful in treating pancreatic cancer  
CC and stimulating an immune response. The polynucleotides can be used as  
CC probes or primers for nucleic acid hybridisation, in the design and  
CC preparation of ribozyme molecules for inhibiting expression of the tumour  
CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
CC therapy. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 473 BP; 129 A; 121 C; 89 G; 134 T; 0 U; 0 Other;  
SQ  
Query Match 96.9%; Score 157; DB 6; Length 473;  
Best Local Similarity 68.2%; Pred. No. 0.8;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
Db |  
QY 61 TTCCATGGACCACTTGGACCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNN 120  
Db |  
QY 121 NNN 157  
Db |  
QY 159 ACCCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 123  
Db |  
RESULT 11  
ACC46770  
ID ACC46770 standard; cdna; 478 BP.  
XX  
XX ACC46770;  
AC  
XX  
XX 05-JUN-2003 (first entry)  
DT  
XX  
XX Human COPD related protein encoding cdna SEQ ID NO:21.  
DE  
XX  
XX Human; chronic obstructive pulmonary disease; COPD; chronic lung disease;  
KW gene; ss.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200297127-A2.  
PN  
XX  
XX 05-DEC-2002.  
PD  
XX  
XX 28-MAY-2002; 2002WO-BP005835.  
PF  
XX  
XX 31-MAY-2001; 2001GB-00013266.  
PR  
XX  
XX (FARB ) BAYER AG.  
PA

PT cancer.  
XX Claim 1; SEQ ID NO 2123; 300pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)  
CC complements of (a); (c) sequences consisting of at least 20 contiguous  
CC residues of (a); (d) sequences that hybridize to (a), under moderately  
CC stringent conditions; (e) sequences having at least 75% or 90% identity  
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
CC ABP68637) encoded by (I) and oligonucleotides can be used to detect cancer  
CC in a patient and compositions comprising polypeptides, polynucleotides,  
CC antibodies, fusion proteins, T cell populations and antigen presenting  
CC cells expressing the polypeptide are useful in treating pancreatic cancer  
CC and stimulating an immune response. The polynucleotides can be used as  
CC probes or primers for nucleic acid hybridisation, in the design and  
CC preparation of ribozyme molecules for inhibiting expression of the tumour  
CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
CC therapy. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 473 BP; 129 A; 121 C; 89 G; 134 T; 0 U; 0 Other;  
SQ  
Query Match 96.9%; Score 157; DB 6; Length 473;  
Best Local Similarity 68.2%; Pred. No. 0.8;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
Db |  
QY 61 TTCCATGGACCACTTGGACCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNN 120  
Db |  
QY 121 NNN 157  
Db |  
QY 159 ACCCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 351  
Db |  
RESULT 10  
ABV97697/c  
ID ABV97697 standard; cdna; 475 BP.  
XX  
XX  
AC  
XX  
XX 14-JAN-2003 (first entry)  
DT  
XX  
XX Human pancreatic cancer expressed cdna SEQ ID NO 3105.  
DE  
XX  
XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;  
KW cytosolic; tumour; gene; ss.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200260317-A2.  
PN  
XX  
XX 08-AUG-2002.  
PD  
XX  
XX 30-JAN-2002; 2002WO-US002781.  
PF  
XX  
XX 30-JAN-2001; 2001US-0265305P.  
PR  
XX  
XX 31-JAN-2001; 2001US-0265682P.  
PR  
XX  
XX 09-FEB-2001; 2001US-0267568P.  
PR  
XX  
XX 21-MAR-2001; 2001US-0278651P.  
PR  
XX  
XX 28-APR-2001; 2001US-0287112P.  
PR  
XX  
XX 16-MAY-2001; 2001US-0291631P.  
PR  
XX  
XX 12-JUL-2001; 2001US-0305484P.  
PR  
XX  
XX 20-AUG-2001; 2001US-0313999P.  
PR  
XX  
XX 27-NOV-2001; 2001US-0333626P.  
PA (CORI-) CORIXA CORP.

```
PI Oellers N, Gehrmann M, Kallabis H, Hall R, Schulze T, Kroegel C;
XX WPI: 2003-140492/13.
DR P-PSDB; ABP96799.
XX
PT Predicting, diagnosing or prognosing chronic lung disease, by detecting a
PT chronic obstructive pulmonary disease (COPD) gene in a biological sample.
XX Claim 8; Page 110; 214pp; English.
XX
CC The present invention describes a method for predicting, diagnosing or
CC prognosing chronic lung disease by detecting a chronic obstructive
CC pulmonary disease (COPD) gene related polynucleotide (see ACC46750 to
CC ACC46777, which encodes the COPD related proteins in ABP96779 to
CC ABP96806). The method is useful for predicting, diagnosing or prognosing
CC chronic lung disease in a biological sample. The COPD genes and proteins
CC encoded by them from the present invention (I) can be used for treating
CC or preventing chronic lung disease in a mammal. (I) can be used in an
CC animal model for determining the efficacy, toxicity, or side effects of
CC treatment with (I), and determining the mechanism of action of (I).
CC ACC46778 to ACC46903 represent COPD related PCR primers and probes used
CC in an example from the present invention
XX
SQ Sequence 478 BP; 142 A; 91 C; 78 G; 167 T; 0 U; 0 Other;
Query Match 96.9%; Score 157; DB 10; Length 478;
Best Local Similarity 68.2%; Pred. No. 0.81;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
DB 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
QY 61 TTCCATGGACCACTTGGACCAAGCAACCAACTCCGAAGACTTCANNNNNNNNNNN 120
DB 61 TTCCATGGACCACTTGGACCAAGCAACCAACTCCGAAGACTTCGAACACCACTCCACA 120
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 121 ACCCAAGACTTCGACCTAATTTATTTCCCTAGCT 157
RESULT 12
ADN95619
ID ADN95619 standard; DNA; 478 BP.
AC ADN95619;
XX
XX 01-JUL-2004 (first entry)
XX Human BBC/LEC-related gene sequence SeqID542.
XX
XX growth; differentiation; blood endothelial cell; BBC;
XX lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
XX lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
XX vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
XX inflammatory disease; cancer metastasis; lymphatic system; gene; ds;
XX human.
XX
XX Homo sapiens.
XX
XX WO2003080640-A1.
XX
XX 02-OCT-2003.
XX
XX 07-MAR-2003; 2003WO-US006900.
XX
XX 07-MAR-2002; 2002US-0363019P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX (LICN-) LICENTIA LTD.
XX
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
```

```
XX WPI: 2003-876899/81.
DR P-PSDB; ADN95618.
XX
PS Example 1; SEQ ID NO 542; 176pp; English.
XX
CC This invention relates to a method of differentially modulating the
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic
CC endothelial cells (LEC) comprises contacting endothelial cells with a
CC composition comprising an agent that differentially modulates blood or
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises
CC identifying a human subject with lymphoedema and with a mutation in at
CC least one allele of a gene encoding a LEC protein, where the mutation
CC correlates with lymphoedema in human subjects, and with the proviso that
CC the LEC protein is not VEGFR-3; and administering to the subject a
CC composition comprising a lymphatic growth agent selected from VEGF-C or
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for
CC the development of compounds with an antiangiogenic, cytostatic,
CC vasotropic or antiinflammatory activity or for gene therapy. The method
CC is useful in modulating the growth or differentiation of blood
CC endothelial cells or lymphatic endothelial cells, in treating hereditary
CC lymphoedema, in screening for an endothelial cell disorder or
CC predisposition to the disorder or in monitoring the efficacy or toxicity
CC of a drug on endothelial cells. The agent is useful in manufacturing a
CC medicament for the differential modulation of blood vessel endothelial
CC cell or lymphatic vessel endothelial cell growth or differentiation. The
CC lymphatic growth agent may also be used in manufacturing a medicament for
CC the treatment of hereditary lymphoedema resulting from a mutation in a
CC LEC gene or of other diseases involving the lymphatic vessels, such as
CC various inflammatory diseases and cancer metastasis via the lymphatic
CC system. The present sequence is that of a human LEC/BEC differentially
CC expressed gene which is related to the method of the invention. Note: This
CC sequence does not appear in the specification but was obtained by the
CC indexer using the source data given in table 14 of the specification.
XX
SQ Sequence 478 BP; 142 A; 91 C; 78 G; 167 T; 0 U; 0 Other;
Query Match 96.9%; Score 157; DB 11; Length 478;
Best Local Similarity 68.2%; Pred. No. 0.81;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
DB 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
QY 61 TTCCATGGACCACTTGGACCAAGCAACCAACTCCGAAGACTTCANNNNNNNNNNN 120
DB 61 TTCCATGGACCACTTGGACCAAGCAACCAACTCCGAAGACTTCGAACACCACTCCACA 120
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 121 ACCCAAGACTTCGACCTAATTTATTTCCCTAGCT 157
RESULT 13
ADP13526
ID ADP13526 standard; DNA; 478 BP.
XX
XX ADP13526;
XX
XX 26-AUG-2004 (first entry)
XX
XX Renal cell carcinoma differentially expressed gene #262.
XX
XX ds; diagnosis; non-blood disease; solid tumor; gene expression;
XX peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
XX head/neck cancer; differential expression.
XX
XX Homo sapiens.
XX
XX WO2004048933-A2.
XX
XX 10-JUN-2004.
```

ss; cytostatic; farnesyl transferase inhibitor; gene expression;  
quinolinone; leukemia; cancer.  
Homo sapiens.  
WO2003038129-A2.  
08-MAY-2003.  
30-OCT-2002; 2002WO-US034784.  
30-OCT-2001; 2001US-0338997P.  
30-OCT-2001; 2001US-0340081P.  
30-OCT-2001; 2001US-0340388P.  
30-OCT-2001; 2001US-0341012P.  
(ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
Raponi M;  
WPI; 2003-513497/48.  
Determining whether a patient will respond to treatment with a farnesyl transferase inhibitor, by analyzing the expression of gene that is differentially modulated in the presence of the inhibitor.  
Disclosure; SEQ ID NO 97; 346pp; English.  
The invention relates to a method of determining whether a patient will respond to treatment with a farnesyl transferase inhibitor (FTI), by analyzing the expression of gene that is differentially modulated in the presence of an FTI. The method is useful for determining whether a patient will respond to treatment with a FTI such as (B)-6-[amino(4-chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-methyl-2-(1H)quinolinone, monitoring the therapy of a patient treating a patient with leukemia with FTI if the analysis indicates that the patient will respond. This sequence corresponds to a gene whose expression may be modulated in the presence of FTI.  
Sequence 482 BP; 143 A; 92 C; 79 G; 168 T; 0 U; 0 Other;  
Query Match 96.9%; Score 157; DB 10; Length 482;  
Best Local Similarity 68.2%; Pred. No. 0.81;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTCAGGA 60  
Db 6 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTCAGGA 65  
QY 61 TTCCATGGACCACTGGACCAAGCAACCCAACTCCGAAGACTTGANNNNNNNNNN 120  
Db 66 TTCCATGGACCACTGGACCAAGCAACCCAACTCCGAAGACTTGAACTCACTCCACA 125  
QY 121 NNN 157  
Db 126 ACCCAAGATCTGCAGCTAACTTATTTTCCCTAGCT 162  
RESULT 15  
ABS68800  
ID ABS68800 standard; RNA; 647 BP.  
XX  
AC ABS68800;  
XX  
XX 20-NOV-2002 (first entry)  
XX Human monocytic chemoattractant protein-1 (hMCP-1) RNA.  
DE  
XX Phosphoramidite; INVADER assay cleavage reaction; PEN1; cleavage;  
KW nucleic acid separation; DNA polymerase; human; MCP-1; ubiquitin;  
KW monocytic chemoattractant protein-1; gene; ss.  
XX  
OS Homo sapiens.

21-NOV-2003; 2003WO-US037481.  
21-NOV-2002; 2002US-0427982P.  
03-APR-2003; 2003US-0459782P.  
(AMHP ) WYETH.  
(TWIN/) TWINE N C.  
(BURC/) BURCZYNSKI M E.  
(TREP/) TREPICCHIO W L.  
(DORN/) DORNER A.  
(STOV/) STOVER J A.  
(SLON/) SLONI D K.  
Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;  
Sloni DK;  
WPI; 2004-460799/43.  
Diagnosing non-blood disease such as solid tumor, involves comparing differential expression profile of specific genes in peripheral blood sample of subject with reference expression profile of specific genes.  
Disclosure; SEQ ID NO 262; 350pp; English.  
The invention relate to a method of diagnosing (M1) non-blood disease such as solid tumor by providing peripheral blood sample of human having non-blood disease, and comparing an expression profile of specific genes in the peripheral blood sample to reference expression profile of the genes, where each of the genes is differentially expressed in peripheral blood mononuclear cells (PBMCs) of patients having the disease as compared to PBMCs of normal humans. The method is useful for diagnosing non-blood disease such as solid tumor. The solid tumor is chosen from renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The peripheral blood sample comprises enriched PBMCs. The peripheral blood sample is a whole blood sample (claimed). (M1) is useful for identifying genes that are differentially expressed in peripheral blood samples isolated at different stages of progression, development or treatment of RCC and/or other solid tumors. This sequence corresponds to a gene that is differentially expressed and detected by the method of the invention. (Note: this sequence is not given as part of the printed specification but was obtained from WIPO in electronic format at ftp.wipo./pub/published\_pct\_sequences).  
Sequence 478 BP; 142 A; 91 C; 78 G; 167 T; 0 U; 0 Other;  
Query Match 96.9%; Score 157; DB 12; Length 478;  
Best Local Similarity 68.2%; Pred. No. 0.81;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTCAGGA 60  
Db 1 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTCAGGA 60  
QY 61 TTCCATGGACCACTGGACCAAGCAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120  
Db 61 TTCCATGGACCACTGGACCAAGCAACCCAACTCCGAAGACTTGAACTCACTCCACA 120  
QY 121 NNN 157  
Db 121 ACCCAAGATCTGCAGCTAACTTATTTTCCCTAGCT 157  
RESULT 14  
ADE84878  
ID ADE84878 standard; DNA; 482 BP.  
XX  
AC ADE84878;  
XX  
XX 29-JAN-2004 (first entry)  
XX Farnesyl transferase inhibitor modulated leukemia associated gene #97.  
DE  
XX



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 16:43:00 ; Search time 1652 Seconds  
(without alignments)  
3732.695 Million cell updates/sec

Title: SEQ13-8698-9073-50N-TGATCA  
Perfect score: 162  
Sequence: 1 cttcaagaccattgtggcca.....nnnnnnnnnnntgatca 162

Scoring table: IDENTITY NUCDX  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	96.9	432	1	AV661469 AV661469
2	157	96.9	446	7	N28510 YX36907.r1
3	157	96.9	486	1	AI478647
4	157	96.9	506	6	CD673147 F521d04.y
5	157	96.9	513	1	AV714555
6	157	96.9	520	4	BG236324 naf26c02
7	157	96.9	541	2	BF115131 hr74g06.x
8	157	96.9	542	7	CK902135 il17e04.x
9	157	96.9	544	2	BF590822 7h40h06.x
10	157	96.9	548	2	AW967017 EST379091
11	157	96.9	549	5	EX490080 DKF2p686B
12	157	96.9	563	4	BM704915 UI-E-C11-
13	157	96.9	565	1	AA614521
14	157	96.9	565	2	BF081639 MR0-AN008
15	157	96.9	569	1	AI892798 wC84b06.x
16	157	96.9	572	1	AV735130
17	157	96.9	581	5	BP197032 BP197032
18	157	96.9	582	1	AA877558 nr05f01.s
19	157	96.9	582	2	BE326715 hr63g10.x
20	157	96.9	586	5	BM997526 UI-H-DH0-
21	157	96.9	588	4	BM708613 UI-E-C11-
22	157	96.9	589	2	AW594110 hg55g02.x
23	157	96.9	593	5	BM888255 TMM174 Hu
24	157	96.9	594	5	BQ631442 il17e04.y

C 25	157	96.9	594	5	BU069097
C 26	157	96.9	598	7	CF370829
C 27	157	96.9	599	7	CK902136
C 28	157	96.9	601	5	BU580378
C 29	157	96.9	601	5	BU580675
C 30	157	96.9	616	1	AL697816
C 31	157	96.9	618	1	AV716988
C 32	157	96.9	621	2	AW273803
C 33	157	96.9	623	2	AW772091
C 34	157	96.9	628	5	BU730951
C 35	157	96.9	634	1	AV713706
C 36	157	96.9	634	5	BQ631169
C 37	157	96.9	635	1	AV717017
C 38	157	96.9	637	1	AV717339
C 39	157	96.9	637	4	BM685551
C 40	157	96.9	640	1	AV734336
C 41	157	96.9	640	5	BM973445
C 42	157	96.9	640	5	BU681561
C 43	157	96.9	640	5	BM685364
C 44	157	96.9	642	5	BM996167
C 45	157	96.9	644	5	BQ573817

ALIGNMENTS

RESULT 1  
AV661469  
LOCUS AV661469 432 bp mRNA linear EST 16-JAN-2002  
DEFINITION AV661469 GLC Homo sapiens cDNA clone GLCGS05 3', mRNA sequence.  
ACCESSION AV661469  
VERSION AV661469.1 GI:9882483  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 432)  
AUTHORS Xu X., Huang J., Xu Z., Qian B., Zhu Z., Yan Q., Cai T., Zhang X., Xiao H., Qu J., Liu F., Huang Q., Cheng Z., Li N., Du J., Hu W., Shen K., Lu G., Fu G., Zhong M., Xu S., Gu W., Huang W., Zhao X., Hu G., Gu J., Chen Z. and Han Z.  
TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
MEDLINE 21625106  
PUBMED 11752456  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..432  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GLCGS05"  
/tissue\_type="corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="GLC"  
/note="Vector: pbluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 96.9%; Score 157; DB 1; Length 432;  
Best Local Similarity 68.2%; Pred. NO. 1.1e-21;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

1	CTTCAAGACCAATTGTGGCCAAAGAGATCTGTCTGACCCCAAGCAGAACTGGGTTTCAGGA	60
93	CTTCAAGACCAATTGTGGCCAAAGAGATCTGTCTGACCCCAAGCAGAACTGGGTTTCAGGA	152
61	TTCCATGGACACCTGGCAAGCAACCAACCAACTCCGAAGACTTCGANNNNNNNNNNNNNN	120
153	TTCCATGGACCACTGGCAAGCAACCAACCAACTCCGAAGACTTGAAGACTTGACACTCCTCCACA	212
121	NN	157
213	ACCCAAGACTCGAGCTAACTTATTTTCCCTTAGCT	249

RESULT 3  
LOCUS AI478647/c  
DEFINITION tm54f08.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2161959 3', similar to gb:M24545 MONOCYTE CHEMOTACTIC PROTEIN 1 PRECURSOR (HUMAN) ;, mRNA sequence.  
ACCESSION AI478647  
VERSION AI478647.1 GI:4371873  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 486)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 975 Std Error: 0.00  
Seq Primer: -40UP from Gibco  
High quality sequence stop: 444.  
Location/Qualifiers  
1. .486  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2161959"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Kid1"  
/note="Organ: kidney; Vector: pRT73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
plasmid DNA from the normalized library NCI CGAP Kid1 was  
prepared, and ss circles were made in vitro. Following HA  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNA  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

```
ORIGIN
Query Match          96.9%; Score 157; DB 1; Length 486;
Best Local Similarity 68.2%; Pred No. 1,2e-21;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0
QY      1   CTTCAAGACATTGTGGCCAAAGAGATCTGTGTCACCCCAAGCAGAGTGGGTTCAGA 60
         |||
DB       482  CTTCAAGACCATTGTGGCCAAAGAGATCTGTGTCACCCCAAGCAGAGTGGGTTCAGA 423
```

1 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTAGTGGGTTCAGA 60  
241 CTTCAAGACCATTGTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAGAGTAGTGGGTTCAGA 300  
61 TTCCATGACACACCTGGGAAGCAAAACCCAACTCCGAAGACTTGGANNNNNNNNNNNN 120  
301 TTCCATGACACACCTGGGAAGCAAAACCCAACTCCGAAGACTTGAACACTCACTCCACA 360  
121 NNN 157  
361 ACCCAAGAACTGTCAGCTAACTTATTTTCCCGTAGCT 397

RESULT 2

N28510 linear EST 04-JAN-1996

LOCUS N28510 446 bp mRNA

DEFINITION Yx36907.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:363868 5' similar to gb:M24545 MONOCYTE CHEMOTACTIC PROTEIN 1 PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION N28510

VERSION N28510.1 GI:1146746

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 446)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Watersson,R., Williamson,A., Wohldmann,P. and

TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 305  
 Source: IMAGE Consortium, LNLNL  
 This clone is available royalty-free through LNLNL ; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: T7  
 High quality sequence stop: 305.  
 Location/Qualifiers  
 i. .446  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3873510"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:263868"  
 /sex="Male"  
 /tissue\_type="melanocyte"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Scorces melanocyte 2NDHM"  
 /note="Vector: p7T73D (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5,  
 TGTTCACCAACTGAAGTGGGCGCGCGACGATTTTTTTTTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified p7T73 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Fátima Bonalido. RNA from normal foreskin melanocytes  
 (FS374) was kindly provided by Dr. Anthony P. Albino."

```

ORIGIN
Query Match          96.9%; Score 157; DB 7; Length 446;
Best Local Similarity 68.2%; Pred. No. 1.1e-21;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

```

QY 61 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120  
 Db 422 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGAACTCACTCCACA 363  
 QY 121 NNT 157  
 Db 362 ACCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 326

RESULT 4  
 CD673147  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 506)  
 Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,  
 Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.  
 TITLE  
 Expressed sequence tag analysis of adult human iris for the NEIBank  
 Project: steroid-response factors and similarities with retinal  
 pigment epithelium  
 JOURNAL  
 Mol. Vis. 8 (4), 185-195 (2002)  
 MEDLINE  
 22103462  
 PUBMED  
 12107412  
 COMMENT  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 21 row: d column: 04  
 Seq primer: M13RPI reverse primer (AB1).  
 Location/Qualifiers  
 1..506  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="fg21d04"  
 /tissue\_type="Iris"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Iris cDNA (Normalized): fg"  
 /note="Organ: Eye; Vector: pCMVSPORT6; A human iris  
 library (bx) was normalized by self-subtraction. One  
 portion of double stranded plasmid DNA representing the  
 library was linearized by NotI. This NotI digested library  
 was used as a template for biotinylated RNA synthesis  
 using SP6 RNA polymerase. Another portion of the double  
 stranded plasmid library was converted to single-stranded  
 circles in vitro using Gene II and Exonuclease III (Life  
 Technologies). Single-stranded DNA (1 mg) was hybridized  
 (Cot 500) with 41 mg of Bio-RNA and vector blocking  
 oligonucleotides. The hybridized Bio-RNA/ss-circles were  
 removed by streptavidin:phenol extraction. EST analysis  
 was performed on the library at the NIH Intramural  
 Sequencing Center (NISC)."

## FEATURES

source  
 1..506  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="fg21d04"  
 /tissue\_type="Iris"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Iris cDNA (Normalized): fg"  
 /note="Organ: Eye; Vector: pCMVSPORT6; A human iris  
 library (bx) was normalized by self-subtraction. One  
 portion of double stranded plasmid DNA representing the  
 library was linearized by NotI. This NotI digested library  
 was used as a template for biotinylated RNA synthesis  
 using SP6 RNA polymerase. Another portion of the double  
 stranded plasmid library was converted to single-stranded  
 circles in vitro using Gene II and Exonuclease III (Life  
 Technologies). Single-stranded DNA (1 mg) was hybridized  
 (Cot 500) with 41 mg of Bio-RNA and vector blocking  
 oligonucleotides. The hybridized Bio-RNA/ss-circles were  
 removed by streptavidin:phenol extraction. EST analysis  
 was performed on the library at the NIH Intramural  
 Sequencing Center (NISC)."

## ORIGIN

Query Match 96.9%; Score 157; DB 6; Length 506;  
 Best Local Similarity 68.2%; Pred. No. 1.2e-21;  
 Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
 Db 130 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 189

QY 61 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120  
 Db 190 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGAACTCACTCCACA 249  
 QY 121 NNT 157  
 Db 250 ACCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 286

RESULT 5  
 AV714555  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 513)  
 Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,  
 Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,  
 Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
 Lu,G., Cheng,Z. and Han,Z.  
 TITLE  
 Homo sapiens cDNA DCB clones  
 JOURNAL  
 Unpublished (2000)  
 COMMENT  
 Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919(ex.45)  
 Fax: 86-21-50801922  
 Email: hanzgchgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 Location/Qualifiers  
 1..513  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DCBADG05"  
 /cell\_type="dendritic cells"  
 /dev\_stage="mature"  
 /lab\_host="BM25.8"  
 /clone\_lib="DCB"  
 /note="Vector: pTriplex2; Site\_1: sf1A; Site\_2: sf1B"

## FEATURES

source  
 1..513  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DCBADG05"  
 /cell\_type="dendritic cells"  
 /dev\_stage="mature"  
 /lab\_host="BM25.8"  
 /clone\_lib="DCB"  
 /note="Vector: pTriplex2; Site\_1: sf1A; Site\_2: sf1B"

## ORIGIN

Query Match 96.9%; Score 157; DB 1; Length 513;  
 Best Local Similarity 68.2%; Pred. No. 1.2e-21;  
 Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
 Db 259 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAGTGGTTTCAGGA 318  
 QY 61 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120  
 Db 319 TTCCATGGACCACTGGCAAGCAAAACCCAACTCCGAAGACTTGAACTCACTCCACA 378  
 QY 121 NNT 157  
 Db 379 ACCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 415

## RESULT 6

BG236324/c  
 LOCUS  
 DEFINITION  
 naf26c02.x1 Soares NPBC Homo sapiens cDNA clone IMAGE:4142186 3'  
 similar to SW:SV02\_HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2  
 PRECURSOR ;, mRNA sequence.  
 ACCESSION  
 BG236324  
 VERSION  
 BG236324.1 GI:12750171

**KEYWORDS**  
SOURCE  
ORGANISM

**REFERENCE**

**AUTHORS**  
**TITLE**

**JOURNAL**  
**COMMENT**

**UNPUBLISHED (1997)**  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-rc@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 425.

**FEATURES**  
source

1..520  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4142186"  
/tissue\_type="lymphocyte"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="Soares APBC"  
/note="Organ: blood; Vector: p7T3D-Pac; Site\_1: NotI; Site\_2: EcoRI; 1st strand cDNA was primed with a Not I - oligo (dt) primer [5', TGTTACCAATCTGAAGTGCGGCGCCGGGTGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library is normalized; constructed in the laboratory of M. Bento Soares (University of Iowa)."

**ORIGIN**

Query Match 96.9%; Score 157; DB 4; Length 520;  
Best Local Similarity 68.2%; Pred. No. 1.2e-21;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTGTGGCCAAGGAGACTGTGCTGACCCCAAGCAGAAGTGGGTTCAGGA 60  
DB 470 CTTCAAGACCATTGTGGCCAAGGAGACTGTGCTGACCCCAAGCAGAAGTGGGTTCAGGA 411

QY 61 TTCCATGGACCACTGGCAAGCAACCAACTCCGAAGACTTTGANNNNNNNNNNNN 120  
DB 410 TTCCATGGACCACTGGCAAGCAACCAACTCCGAAGACTTTGAACTCATCTCCACA 351

QY 121 NNN 157  
DB 350 ACCAAGATCTGCAGCTTAACCTATTATTTCCCCTAGCT 314

**RESULT 7**  
BF115131/c  
LOCUS  
DEFINITION

BF115131 541 bp mRNA linear EST 24-OCT-2000  
hr74g06.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:3134266 3'  
similar to SW\_Sy02\_HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2  
PRECURSOR ;, mRNA sequence.

**ACCESSION**  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BF115131 GI:10984607  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**NCI-CGAP** http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-rc@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 471.

**FEATURES**  
source

1..541  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3134266"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Kid11"  
/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid1 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified CDNA from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

**ORIGIN**

Query Match 96.9%; Score 157; DB 2; Length 541;  
Best Local Similarity 68.2%; Pred. No. 1.3e-21;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTGTGGCCAAGGAGACTGTGCTGACCCCAAGCAGAAGTGGGTTCAGGA 60  
DB 479 CTTCAAGACCATTGTGGCCAAGGAGACTGTGCTGACCCCAAGCAGAAGTGGGTTCAGGA 420

QY 61 TTCCATGGACCACTGGCAAGCAACCAACTCCGAAGACTTTGANNNNNNNNNNNN 120  
DB 419 TTCCATGGACCACTGGCAAGCAACCAACTCCGAAGACTTTGAACACTCATCTCCACA 360

QY 121 NNN 157  
DB 359 ACCCAAGAATCTGCAGCTTAACCTATTATTTCCCCTAGCT 323

**RESULT 8**  
CK902135  
LOCUS  
DEFINITION

CK902135 542 bp mRNA linear EST 11-MAR-2004  
1117e04.x6 HR85 islet Homo sapiens cDNA clone IMAGE:6030414 3'  
similar to SW\_Sy02\_HUMAN P13500 SMALL INDUCIBLE CYTOKINE A2  
PRECURSOR ;, mRNA sequence.

**ACCESSION**  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CK902135 GI:45363666  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**  
1 (bases 1 to 542)  
Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
WashU-Harvard Pancreas EST Project



```

Seq primer: Reverse.
Location/Qualifiers
1..548
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGJ"
/note="Vector: pBluescriptSKm"

ORIGIN
Query Match          96.9%; Score 157; DB 2; Length 548;
Best Local Similarity 68.2%; Pred. No. 1.3e-21;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
Db 263 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 322

QY 61 TTCCATGGACCACTGGACCAAGCAAAACCCAAACTCCGAAGACTTGGANNNNNNNNNNN 120
Db 323 TTCCATGGACCACTGGACCAAGCAAAACCCAAACTCCGAAGACTTGAACACTCACTCCACA 382

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db 383 ACCCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 419

RESULT 12
LOCUS BM704915
DEFINITION UI-E-C11-aggf-h-04-0-UI.r1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION BM704915
VERSION BM704915.1 GI:19018173
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 563)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="UI-E-C11-aggf-h-04-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
sequence tag for this library is ACTCA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

FEATURES
source
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wienann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wienann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFP686B1371) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="DKFP686B1371"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site 1: SfiIA; Site 2: SfiIB;
cDNA-collection"

ORIGIN
Query Match          96.9%; Score 157; DB 5; Length 549;
Best Local Similarity 68.2%; Pred. No. 1.3e-21;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60
Db 238 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 297

QY 61 TTCCATGGACCACTGGACCAAGCAAAACCCAAACTCCGAAGACTTGGANNNNNNNNNNN 120
Db 298 TTCCATGGACCACTGGACCAAGCAAAACCCAAACTCCGAAGACTTGAACACTCACTCCACA 357

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
Db 358 ACCCAAGAATCTGCAGCTAACTATTATTTCCCTAGCT 394

RESULT 11
LOCUS BX490080
DEFINITION BX490080 549 bp mRNA linear EST 04-SEP-2003
ACCESSION DKFP686B1371_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
VERSION DKFP686B1371 5', mRNA sequence.
KEYWORDS EST.
SOURCE BX490080.1 GI:31999304
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 549)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbauer,S., Mewes,H.W.,
Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Ottenwaelder,B., Obermaier,B., Deutschenbauer,S., Mewes,H.W.,
et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wienann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wienann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFP686B1371) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="DKFP686B1371"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site 1: SfiIA; Site 2: SfiIB;
cDNA-collection"

ORIGIN
Query Match          96.9%; Score 157; DB 5; Length 549;
Best Local Similarity 68.2%; Pred. No. 1.3e-21;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

```

## ORIGIN

Query Match 96.9%; Score 157; DB 4; Length 563;  
Best Local Similarity 68.2%; Pred. No. 1.3e-21;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTGCTGACCCCAAGCAGAAAGTGGTTTCAGGA 60  
|||||  
DB 233 CTTCAAGACCATTTGGCCCAAGAGATCTGTGCTGACCCCAAGCAGAAAGTGGTTTCAGGA 292  
|||||

QY 61 TTCCATGGACCACTGGCAAGCAACCAACCTCGAAGACTTGANNNNNNNNNNN 120  
|||||  
DB 293 TTCCATGGACCACTGGCAAGCAACCAACCTCGAAGACTTGAACACTCACTCCACA 352  
|||||

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157  
|||||  
DB 353 ACCCAAGATCTGCAGCTAACTTATTTTCCCTAGCT 389  
|||||

RESULT 13  
AA614521/c

LOCUS  
DEFINITION  
AA614521.1 GI:2466717  
565 bp mRNA linear EST 16-OCT-1997  
np49b05.e1 NCI CGAP Br1.1 Homo sapiens cDNA clone IMAGE:1129617 3'  
similar to gb:M24545 MONOCYTE CHEMOTACTIC PROTEIN 1 PRECURSOR  
(HUMAN);, mRNA sequence.

ACCESSION  
VERSION  
AA614521  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 565)

REFERENCE  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 811 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 463.  
Location/Qualifiers  
1. .565  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1129617"  
/sex="female, pooled"  
/tissue\_type="breast"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Br1.1"  
/note="Vector: pVT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from pooled bulk  
breast tumor tissue, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pVT73  
vector. Library is not normalized. (The normalized  
version of this library is NCI CGAP Br2.) Library was  
constructed by Bento Soares and M. Fatima Bonaldo."

## FEATURES

## source

FEATURES  
source

1. .565  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="AN0084"  
/note="Organ: amnion normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 96.9%; Score 157; DB 2; Length 565;  
Best Local Similarity 68.8%; Pred. No. 1.3e-21;  
Matches 108; Conservative 49; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTGCTGACCCCAAGCAGAAAGTGGTTTCAGGA 60

Job time : 1657 secs

Db 263 CTTCAAGACCATTTGGCCAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 322  
QY 61 TTCCATGGACCACTTGGCAAGCAAAACCAACTCCGAAGACTTGANNNNNNNNNNN 120  
Db 323 TTCCATGGACCACTTGGCAAGCAAAACCAACTCCGAAGACTTGAACATTCCTCCACA 382  
QY 121 NNN 157  
Db 383 ANCCAAGAACTGCAGCTAACTATTTCCTCCCTAGCT 419

RESULT 15  
AI692798/c  
LOCUS AI692798 569 bp mRNA linear EST 17-DEC-1999  
DEFINITION wc84b06.x1 NCI CGAP Pan1 Homo sapiens cDNA clone IMAGE:2325299 3',  
similar to gb:M24545 MONOCYTE CHEMOTACTIC PROTEIN 1 PRECURSOR  
(HUMAN);, mRNA sequence.  
ACCESSION AI692798  
VERSION AI692798.1 GI:4970138  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 569)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert length: 707 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 417.

FEATURES  
source  
1..569  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2325299"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/clone\_lib="NCI-CGAP\_Pan1"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

ORIGIN  
Query Match 96.9%; Score 157; DB 1; Length 569;  
Best Local Similarity 68.2%; Pred. No. 1.3e-21;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTCAAGACCATTTGGCCAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 60  
Db 478 CTTCAAGACCATTTGGCCAGGAGATCTGTGCTGACCCCAAGCAGAGTGGTTTCAGGA 419  
QY 61 TTCCATGGACCACTTGGCAAGCAAAACCAACTCCGAAGACTTGANNNNNNNNNNN 120  
Db 418 TTCCATGGACCACTTGGCAAGCAAAACCAACTCCGAAGACTTGAACATTCCTCCACA 359  
QY 121 NNN 157  
Db 358 ACCCAAGAACTGCAGCTAACTATTTCCTCCCTAGCT 322

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 16:44:28 ; Search time 83 seconds  
(without alignments)  
3193.697 Million cell updates/sec

Title: SEQ13-8698-9073-50N-TGATCA  
Perfect score: 162  
Sequence: 1 cttcaagaccattgtgccca.....nnnnnnnnnnntgatca 162

Scoring table: IDENTITY\_NUCDX  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	96.9	647	4	US-09-777-430C-72
2	157	96.9	661	4	US-09-814-915A-104
3	157	96.9	725	4	US-09-591-992-3
4	157	96.9	725	4	US-09-023-655-1344
5	157	96.9	741	3	US-07-927-391-25
6	157	96.9	741	4	US-09-023-655-1165
7	157	96.9	772	4	US-08-437-306-1
8	157	96.9	1712	4	US-09-148-545-106
9	157	96.9	1822	4	US-09-148-545-105
10	140	86.4	752	6	5212073-1
11	140	86.4	752	6	5212073-1
12	132.4	81.7	810	4	US-09-016-434-1273
13	132.4	81.7	810	4	US-09-949-016-3840
14	132.4	81.7	814	3	US-07-927-391-15
15	132.4	81.7	6022	4	US-09-949-016-15582
16	124.2	76.7	353	3	US-09-366-887A-21
17	124.2	76.7	353	4	US-09-517-204-21
18	124.2	76.7	605	3	US-09-366-887A-26
19	124.2	76.7	605	4	US-09-517-204-26
20	124.2	76.7	807	4	US-09-023-655-978
21	124.2	76.7	4119	4	US-09-076-259-1
22	122.6	75.7	813	4	US-09-016-434-1156
23	122.6	75.7	994	3	US-09-366-887A-5
24	122.6	75.7	994	4	US-09-517-204-5
25	119.4	73.7	802	4	US-09-016-434-687
26	119.4	73.7	823	4	US-09-545-894-1
27	115.4	71.2	840	4	US-09-016-434-1033

28	112.8	69.6	514	4	US-09-545-894-3	Sequence 3, Appli
29	112.8	69.6	540	3	US-08-744-419-1	Sequence 1, Appli
30	109.2	67.4	818	3	US-09-366-887A-15	Sequence 15, Appli
31	109.2	67.4	818	4	US-09-517-204-15	Sequence 15, Appli
32	109	67.3	213	4	US-09-463-451-29	Sequence 29, Appli
33	109	67.3	213	4	US-09-463-451-30	Sequence 30, Appli
34	108.2	66.8	1086	4	US-09-646-028-37	Sequence 37, Appli
35	107.4	66.3	1080	4	US-09-646-028-39	Sequence 39, Appli
36	106.8	65.9	1805	4	US-09-646-028-45	Sequence 45, Appli
37	105.6	65.2	1113	4	US-09-646-028-38	Sequence 38, Appli
38	105.4	65.1	1047	4	US-09-646-028-44	Sequence 44, Appli
39	101.4	62.6	231	5	PCT-US95-00605-2	Sequence 2, Appli
40	99.8	61.6	228	1	US-08-250-958-3	Sequence 3, Appli
41	90.4	55.8	360	2	US-08-479-126B-1	Sequence 1, Appli
42	90.4	55.8	360	2	US-08-726-830A-1	Sequence 1, Appli
43	90.4	55.8	360	3	US-08-995-156A-1	Sequence 1, Appli
44	90.4	55.8	360	3	US-09-044-856A-1	Sequence 1, Appli
45	90.4	55.8	360	3	US-09-044-855A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-777-430C-72  
; Sequence 72, Application US/09777430C  
; Patent No. 6780982  
; GENERAL INFORMATION:  
; APPLICANT: Lyamichev, Victor  
; APPLICANT: Skrzypczynski, Zbigniew  
; APPLICANT: Allawi, Hatim T.  
; APPLICANT: Wayland, Sarah R.  
; APPLICANT: Takova, Tsatska  
; APPLICANT: Neir, Bruce P.  
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules  
; FILE REFERENCE: FORS-04912  
; CURRENT APPLICATION NUMBER: US/09/777,430C  
; CURRENT FILING DATE: 2001-02-06  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 72  
; LENGTH: 647  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-09-777-430C-72

Query Match	96.9%	Score 157;	DB 4;	Length 647;
Best Local Similarity	55.4%	Pred. NO. 0.014;		
Matches	87;	Conservative	70;	Mismatches 0; Indels 0; Gaps 0;
QY	1	CTTCAAGACCATTTGGCCCAAGGAGATCTGTGTCACCCCAAGCAGAGAGTGGGTTTCAGGA	60	
DB	195	CTUCAAGACCAUUGTGGCCCAAGGAGCUGUCUGACCCCAAGCAGAGAGGUGGUUCAGGA	254	
QY	61	TTTCATGACCACTTGGCAAGCAACCAAACTCCGAAGACTTGCANNNNNNNNNNN	120	
DB	255	UUCUAGGACCAACCCGACCAAGCAACCAACUCCGAGACUUGACACUCACUCCACCA	314	
QY	121	NN	157	
DB	315	ACCCAAGAAUCUGCAGCUACUAUUUCCCUCCUAGCU	351	

RESULT 2  
US-09-814-915A-104  
; Sequence 104, Application US/09814915A  
; Patent No. 6750015  
; GENERAL INFORMATION:  
; APPLICANT: Horwitz, Kathryn  
; APPLICANT: Richer, Jennifer  
; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat

```
; TITLE OF INVENTION: Thereto
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/09/814,915A
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; LENGTH: 661
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-814-915A-104

Query Match          96.9%; Score 157; DB 4; Length 661;
Best Local Similarity 68.2%; Pred. No. 0.015;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 60
DB 226 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 285

QY 61 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGANNNNNNNNNNNN 120
DB 286 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGAAACACTCCTCCACA 345

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 346 ACCCAAGAATCTGCAGCTAACTTAATTTCCCTAGCT 382

RESULT 3
US-09-591-992-3
; Sequence 3, Application US/09591992
; Patent No. 6569418
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Devico, Anthony L.
; APPLICANT: Garzino, Alfredo
; TITLE OF INVENTION: Immuno-Modulating Effects of Chemokines in DNA Vaccination
; FILE REFERENCE: 4115-109 CIP
; CURRENT APPLICATION NUMBER: US/09/591,992
; CURRENT FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PCT/US98/26291
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/186,416
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/069,281
; PRIOR FILING DATE: 1997-12-11
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-591-992-3

Query Match          96.9%; Score 157; DB 4; Length 725;
Best Local Similarity 68.2%; Pred. No. 0.015;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 60
DB 248 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 307

QY 61 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGANNNNNNNNNNNN 120
DB 308 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGAAACACTCCTCCACA 367

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 368 ACCCAAGAATCTGCAGCTAACTTAATTTCCCTAGCT 404
```

```
RESULT 4
US-09-023-655-1344
; Sequence 1344, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1344:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 725 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 934513
; US-09-023-655-1344

Query Match          96.9%; Score 157; DB 4; Length 725;
Best Local Similarity 68.2%; Pred. No. 0.015;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 60
DB 248 CTTCAAGACCAATTGTGGCCCAAGGAGATCTGTCTGACCCCAAGCAGAAAGTGGTTTCAGGA 307

QY 61 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGANNNNNNNNNNNN 120
DB 308 TTCCATGGACCACTGGGACCAAGCAACCAACCAACTCGGAGACTTCGAAACACTCCTCCACA 367

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 368 ACCCAAGAATCTGCAGCTAACTTAATTTCCCTAGCT 404

RESULT 5
US-07-927-391-25
; Sequence 25, Application US/07927391
; Patent No. 6001649
; GENERAL INFORMATION:
; APPLICANT: CAPUT, Daniel
```

APPLICANT: FERRARA, Pascual  
APPLICANT: MILOUX, Brigitte  
APPLICANT: MINTY, Adrian  
APPLICANT: VITA, Natalio  
TITLE OF INVENTION: Protein having a cytokin type  
TITLE OF INVENTION: activity, and recombinant DNA, expression vector and hosts  
TITLE OF INVENTION: for its preparation.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: King Street Station, Suite 500, 1800 Diagonal  
STREET: Road, PO Box 239  
CITY: ALEXANDRIA  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/927,391  
FILING DATE: 19920929  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Bernhard D  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 16781/369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 741 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-927-391-25

Query Match 96.9%; Score 157; DB 3; Length 741;  
Best Local Similarity 68.2%; Pred. No. 0.015;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60  
DB 264 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 323

QY 61 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGANNNNNNNNNNNN 120  
DB 324 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGAACTCACTCCACA 383

QY 121 NNN 157  
DB 384 ACCCAAGATCTGCAGCTAACTATTATTTCCCTAGCT 420

RESULT 6  
US-09-023-655-1165  
Sequence 1165, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE

APPLICANT: FERRARA, Pascual  
APPLICANT: MILOUX, Brigitte  
APPLICANT: MINTY, Adrian  
APPLICANT: VITA, Natalio  
TITLE OF INVENTION: Protein having a cytokin type  
TITLE OF INVENTION: activity, and recombinant DNA, expression vector and hosts  
TITLE OF INVENTION: for its preparation.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: King Street Station, Suite 500, 1800 Diagonal  
STREET: Road, PO Box 239  
CITY: ALEXANDRIA  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/927,391  
FILING DATE: 19920929  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Saxe, Bernhard D  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 16781/369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 741 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-07-927-391-25

Query Match 96.9%; Score 157; DB 3; Length 741;  
Best Local Similarity 68.2%; Pred. No. 0.015;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60  
DB 264 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 323

QY 61 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGANNNNNNNNNNNN 120  
DB 324 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGAACTCACTCCACA 383

QY 121 NNN 157  
DB 384 ACCCAAGATCTGCAGCTAACTATTATTTCCCTAGCT 420

RESULT 6  
US-09-023-655-1165  
Sequence 1165, Application US/09023655  
Patent No. 6607879  
GENERAL INFORMATION:  
APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan G. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE

CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1165:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 741 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g187434  
US-09-023-655-1165

Query Match 96.9%; Score 157; DB 4; Length 741;  
Best Local Similarity 68.2%; Pred. No. 0.015;  
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60  
DB 264 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 323

QY 61 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGANNNNNNNNNNNN 120  
DB 324 TTCCATGGACCACTGGCAAGCAAGCAACCAAACTCCGAAGACTTCGAACTCACTCCACA 383

QY 121 NNN 157  
DB 384 ACCCAAGATCTGCAGCTAACTATTATTTCCCTAGCT 420

RESULT 7  
US-08-437-306-1  
Sequence 1, Application US/08437306  
Patent No. 6787645  
GENERAL INFORMATION:  
APPLICANT: Rollins, Barrett J.  
APPLICANT: Stiles, Charles D.  
APPLICANT: Wong, Gordon G.  
TITLE OF INVENTION: No. 6787645el Human Cytokine  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

Tue Aug 9 17:01:23 2005

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,306
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/228,931
; FILING DATE: 13-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/003,136
; FILING DATE: 12-JAN-1993
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,515
; FILING DATE: 16-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/351,008
; FILING DATE: 12-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: DFCI-196A2Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 772 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..369
; US-08-437-306-1

Query Match 96.9%; Score 157; DB 4; Length 772;
Best Local Similarity 68.2%; Pred. No. 0.015;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTGTGACCCCAAGCAGAACTGGGTTTCAGGA 60
DB 267 CTTCAAGACCATTTGGCCCAAGAGATCTGTGTGACCCCAAGCAGAACTGGGTTTCAGGA 326
QY 61 TTCCATGGACCACTGGCAGCAGCAAAACCCAACTCCGAACTTGGANNNNNNNNNNN 120
DB 327 TTCCATGGACCACTGGCAGCAGCAAAACCCAACTCCGAACTTGAACACTCACTCCACA 386
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 387 ACCCAAGAATCGCAGCTAACTTTTCCCTTAGCT 423

RESULT 8
US-09-148-545-106
; Sequence 106, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001PI
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
```

```

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 05-Sep-1997
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23

```

```

; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 1712

Query Match          96.9%; Score 157; DB 4; Length 1712;
Best Local Similarity 68.2%; Pred No. 0.025;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAAGTGGTTCAGCA 60
    |||||
Db 1184 CTTCAAGACCATTTGGCCAAAGGAGATCTGTGCTGACCCCAAGCAAGTGGTTCAGCA 1243

QY 61 TTCCATGCAACCACTGGGACCAAGCAACCAAACTCCGAAGACTTGANNNNNNNNNNN 120
    |||||
Db 1244 TTCCATGCAACCACTGGGACCAAGCAACCAAACTCCGAAGACTTGAACACTCACTCCACA 1303

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 157
    |||||
Db 1304 ACCCAAGATCTGCAGCTAATTTTCCCTAGCT 1340

RESULT 9
US-09-148-545-105
; Sequence 105, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: PZ001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161

```

[illegible]



```
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1273:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: q288396
; US-09-016-434-1273

Query Match      81.7%; Score 132.4; DB 4; Length 810;
Best Local Similarity 58.2%; Pred. No. 1.3;
Matches 92; Conservative 50; Mismatches 16; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
DB 265 CTTCAAGACCAACTGGACAGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 324

QY 61 TTCCATGGACCACTGGACAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 325 CTTTATGAAGCACTGGACAGAGAAACCCAACTCCAAAGCTTTGAACATTTCATGACTGA 384

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTG 158
DB 385 ACTAAAAACAAGCCATGACTTGTGAGAAACAATAATTG 422

RESULT 13
US-09-949-016-3840
; Sequence 3840, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3840
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-3840

Query Match      81.7%; Score 132.4; DB 4; Length 810;
Best Local Similarity 58.2%; Pred. No. 1.3;
Matches 92; Conservative 50; Mismatches 16; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
DB 265 CTTCAAGACCAACTGGACAGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 324

QY 61 TTCCATGGACCACTGGACAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 325 CTTTATGAAGCACTGGACAGAGAAACCCAACTCCAAAGCTTTGAACATTTCATGACTGA 384

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTG 158
DB 385 ACTAAAAACAAGCCATGACTTGTGAGAAACAATAATTG 422

Db 385 ACTGAAAAACAAGCCATGACTTGTGAGAAACAATAATTG 422

RESULT 14
US-07-927-391-15
; Sequence 15, Application US/07927391
; Patent No. 6001649
; GENERAL INFORMATION:
; APPLICANT: CAPUT, Daniel
; APPLICANT: FERRARA, Pascual
; APPLICANT: MILOUX, Brigitte
; APPLICANT: MINTY, Adrian
; APPLICANT: VITA, Natalio
; TITLE OF INVENTION: Protein having a cytokin type
; TITLE OF INVENTION: activity, and recombinant DNA, expression vector and hosts
; TITLE OF INVENTION: for its preparation.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: King Street Station, Suite 500, 1800 Diagonal
; STREET: Road, PO Box 299
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/927,391
; FILING DATE: 19920929
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 16781/369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 814 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41...367
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 140...367
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 41...139
; US-07-927-391-15

Query Match      81.7%; Score 132.4; DB 3; Length 814;
Best Local Similarity 58.2%; Pred. No. 1.3;
Matches 92; Conservative 50; Mismatches 16; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 60
DB 265 CTTCAAGACCAACTGGACAGAGATCTGTGCTGACCCCAAGCAGAGTGGGTTTCAGGA 324

QY 61 TTCCATGGACCACTGGACAGCAAAACCCAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 325 CTTTATGAAGCACTGGACAGAGAAACCCAACTCCAAAGCTTTGAACATTTCATGACTGA 384
```

Qy 121 NNTG 158  
Db 385 ACTAAACACAGCCATGACTTGAGAAACAAATAATTG 422

## RESULT 15

US-09-949-016-15582  
; Sequence 15582, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15582  
; LENGTH: 6022  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15582

Query Match 81.7%; Score 132.4; DB 4; Length 6022;  
Best Local Similarity 58.2%; Pred. No. 4.6;  
Matches 92; Conservative 50; Mismatches 16; Indels 0; Gaps 0;  
Qy 1 CTTCAAGACCATTTGGCCCAAGGAGATCTGTGACCCCAAGCAGAGTGGGTTCCAGGA 60  
Db 3477 CTTCAAGACCAAACTGGACCAAGGAGATCTGTGACCCCAAGGAGTGGGTTCCAGGA 3536  
Qy 61 TTCCATGGACCACTGGCAAGCAACCAAACTCCGAAGACTTGANNNNNNNNNNNN 120  
Db 3537 CTTTATGAGACCTGGACAGAAACCCAACTTCAAGCTTTGAACATTCATGACTGA 3596  
Qy 121 NNTG 158  
Db 3597 ACTGAAACACAGCCATGACTTGAGAAACAAATAATTG 3634

Search completed: August 4, 2005, 17:40:05  
Job time : 84 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 16:50:19 ; Search time 332 Seconds  
(without alignments)  
3163.062 Million cell updates/sec

Title: SEQ13-8698-9073-50N-TGATCA

Perfect score: 162

Sequence: 1 cttaagaccattgtggcca.....nnnnnnnnnnntgatca 162

Scoring table: IDENTITY\_NUCDX

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*

13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*

14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*

15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*

16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*

17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*

18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*

19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*

20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*

21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*

22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*

23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*

24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*

25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	157	96.9	360	9	US-09-864-761-5454
2	157	96.9	473	14	US-10-060-036-2123
3	157	96.9	475	14	US-10-060-036-3105
4	157	96.9	478	19	US-10-717-597-262
5	157	96.9	482	19	US-10-283-975A-97
6	157	96.9	647	9	US-09-777-430A-72
7	157	96.9	647	22	US-10-875-094-72

RESULT 1

US-09-864-761-5454/c  
; Sequence 5454, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30

ALIGNMENTS

c

Sequence 104, Appl  
Sequence 66, Appl  
Sequence 1344, Ap  
Sequence 58, Appl  
Sequence 17, Appl  
Sequence 3, Appl  
Sequence 46, Appl  
Sequence 54, Appl  
Sequence 66, Appl  
Sequence 395, App  
Sequence 849, App  
Sequence 849, App  
Sequence 96, Appl  
Sequence 35, Appl  
Sequence 549, App  
Sequence 1165, Ap  
Sequence 2, Appli  
Sequence 210, App  
Sequence 1, Appli  
Sequence 377, App  
Sequence 1, Appli  
Sequence 516, App  
Sequence 209, App  
Sequence 18, Appl  
Sequence 583, App  
Sequence 211, App  
Sequence 9822, Ap  
Sequence 41, Appl  
Sequence 106, App  
Sequence 105, App  
Sequence 105, App  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 76, Appl  
Sequence 54, Appl  
Sequence 54, Appl  
Sequence 52, Appl

```
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5454
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005549.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; US-08-864-761-5454

Query Match          96.9%; Score 157; DB 9; Length 360;
Best Local Similarity 68.2%; Pred. No. 1.7;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60
DB 358 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 299

QY 61 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 298 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGAAACACTCACTCCACA 239

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 238 ACCCAAGAATCTGCAGCTAACTTATTTTCCCTAGCT 202

RESULT 2
US-10-060-036-2123
; Sequence 2123, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5454
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-060-036-2123

Query Match          96.9%; Score 157; DB 14; Length 475;
Best Local Similarity 68.2%; Pred. No. 1.8;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60
DB 279 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 220

QY 61 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 219 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGAAACACTCACTCCACA 160

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 159 ACCCAAGAATCTGCAGCTAACTTATTTTCCCTAGCT 123

RESULT 3
US-10-060-036-3105/c
; Sequence 3105, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3105
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-060-036-3105

Query Match          96.9%; Score 157; DB 14; Length 475;
Best Local Similarity 68.2%; Pred. No. 1.8;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 60
DB 279 CTTCAAGACCATTTGGCCCAAGAGAGATCTGTGCTGACCCCAAGCAGAGAGTGGGTTTCAGGA 220

QY 61 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
DB 219 TTCCATGGACCACTTGGCAAGCAAGCAAACTCCGAAGACTTGAAACACTCACTCCACA 160

QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
DB 159 ACCCAAGAATCTGCAGCTAACTTATTTTCCCTAGCT 123

RESULT 4
US-10-717-597-262
; Sequence 262, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
```

```
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 262
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-717-597-262

Query Match          96.9%; Score 157; DB 19; Length 482;
Best Local Similarity 68.2%; Pred. No. 1.8;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 60
Db 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 60
QY 61 TTCCATGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 61 TTCCATGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGAAACACCACTCCACA 120
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 121 ACCCAAGATCTGCAGCTAACTATTATTTCCCTAGCT 157

RESULT 5
US-10-283-975A-97
; Sequence 97, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 482
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-97

Query Match          96.9%; Score 157; DB 19; Length 482;
Best Local Similarity 68.2%; Pred. No. 1.8;
Matches 107; Conservative 50; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 60
Db 6 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 65
QY 61 TTCCATGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 66 TTCCATGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGAAACACTCACTCCACA 125
```

```
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 126 ACCCAAGATCTGCAGCTAACTATTATTTCCCTAGCT 162

RESULT 6
US-09-777-430A-72
; Sequence 72, Application US/09777430A
; Patent No. US20020128465A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Skrzypczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Tsetska
; APPLICANT: Neir, Bruce P.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: FORS-04912
; CURRENT APPLICATION NUMBER: US/09/777,430A
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 647
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-777-430A-72

Query Match          96.9%; Score 157; DB 9; Length 647;
Best Local Similarity 55.4%; Pred. No. 1.9;
Matches 87; Conservative 70; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCAAGACCATTTGGCCCAAGAGATCTGTCTGACCCCAAGCAGAAAGTGGGTTTCAGGA 60
Db 195 CUUCAAGACCAUUGUGGCCCAAGGAGUCUGUGUGACCCCAAGCAGAAAGUGGUGUUCAGGA 254
QY 61 TTCCATGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGANNNNNNNNNNNN 120
Db 255 UUCAAGGACCACTGGGCAAGCAACCAAACTCCGAAGACTTGAAACACACUCACUCCACA 314
QY 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157
Db 315 ACCCAAGAAUCUGCAGCUAACUUAUUUCCCUAGCU 351

RESULT 7
US-10-875-094-72
; Sequence 72, Application US/10875094
; Publication No. US20050130179A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Skrzypczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Tsetska
; APPLICANT: Neir, Bruce P.
; APPLICANT: Prudent, James R.
; APPLICANT: Hall, Jeff G.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: FORS-04912
; CURRENT APPLICATION NUMBER: US/10/875,094
; CURRENT FILING DATE: 2004-06-23
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 647
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
```





[illegible]

## RESULT 14

```

US-10-852-335A-46
; Sequence 46, Application US/10852335A
; Publication No. US20050112129A1
; GENERAL INFORMATION:
; APPLICANT: HEIDI S. PHILLIPS
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; Treatment of Tumors of Glial Origin
; FILE REFERENCE: P5103R1-US
; CURRENT APPLICATION NUMBER: US/10/852,335A
; CURRENT FILING DATE: 2004-05-24
; PRIOR APPLICATION NUMBER: US 60/548,299
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/473,238
; PRIOR FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 46
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-852-335A-46

```

## RESULT 15

```

US-10-852-335A-54
; Sequence 54, Application US/10852335A
; Publication No. US20050112129A1
; GENERAL INFORMATION:
; APPLICANT: HEIDI S. PHILLIPS
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITL OF INVENTION: Treatment of Tumors of Glial Origin
; FILE REFERENCE: P5103R1-US
; CURRENT APPLICATION NUMBER: US/10/852,335A
; CURRENT FILING DATE: 2004-05-24
; PRIOR APPLICATION NUMBER: US 60/548,299
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/473,238
; PRIOR FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 190
; SEQ ID NO 54
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-852-335A-54

```

Db	248	CTTCAAGACCATGTGTGGCCCAAGGATCTGTGCTGACCCCAAGCAGAAAGTGGGTT	CAGGA	307
Qy	61	TTCCATGGACCACTTGGGACAGACAAACCCAAATCTCCGAGACTTGANNNNNNNNNNNN		120
Db	308	TTCCATGGACCACTTGGGACAGACAAACCCAAATCTCCGAGACTTGAACACTCACTCCACA		367
Qy	121	NN		157
Db	368	ACCCAGAAGATCTGCAGCTAACTTATTTTCCCTCCAGCT		404

Search completed: August 4, 2005, 17:45:48  
Job time : 334 secs